*Proc. Nat. Acad. Sci. USA*
Vol. 70, No. 6, pp. 1819–1823, June 1973

# Moments of the End-to-End Vector of a Chain Molecule, Its Persistence and Distribution

PAUL J. FLORY

Department of Chemistry, Stanford University, Stanford, California 94305

*Contributed by Paul J. Flory, March 19, 1973*

**ABSTRACT** **The persistence vector a is defined as the configurational average of the chain vector r connecting the ends of the molecule and expressed in a reference frame fixed with respect to the first two skeletal bonds. Moments of second and higher orders in the components of r may readily be calculated for real chains in the rotational isomeric state approximation, and from them the corresponding moments of the vector $\rho = \mathbf{r} - \mathbf{a}$ measured from the terminus of a. Development of the density distribution of $\rho$ about a is proposed as an alternative to the customary treatment of the density distribution of r about $r = 0$ on the assumption that this latter distribution should be (approximately) symmetric. Past difficulties in the analysis of cyclization equilibria involving rings of moderate size, such as occur in single strands of polynucleotide chains, conceivably may be overcome by adoption of this alternative.**

This paper is concerned with statistical mechanical averages of components of the vector **r** connecting the ends of a chain molecule and with averages of products of these components. If the vector **r** is expressed in a reference frame attached to the molecule, moments involving odd as well as even powers of the vector components may take on non-zero values, averaging being performed over all internal configurations of the chain. These moments present a much fuller array of information on the character of the chain configuration than is offered by the even moments $\langle r^2\rangle$, $\langle r^4\rangle$, etc. of the magnitude of **r**. Attention heretofore has been devoted almost exclusively to these latter quantities (1–4). The moments here treated are especially important for the analysis of characteristics of chains of short or intermediate lengths.

Averages of the components of **r** when expressed in a coordinate system affixed to the first two bonds of the chain define a *persistence vector* **a**. This vector for a real chain is the analog of the persistence *length* of the familiar Porod–Kratky model chain (5). Its direction and magnitude depend on the geometry and configurational characteristics peculiar to the given chain. Unlike the Porod–Kratky persistence length, it will not, in general, be directed along the initial bond, or element, of the chain. The components of the persistance vector **a** (i.e., the first moments of the components of **r**) in conjunction with the matrix of second moments specify the density distribution function of the vector **r** in Gaussian approximation relative to the mean displacement **a** from the beginning of the sequence of units, which we designate as the chain. Higher moments serve to define this distribution in higher approximation.

The various moments may be evaluated readily by adaptation of matrix multiplication methods developed in recent years (3, 6–8).

## MOMENTS OF FIRST ORDER; THE PERSISTENCE VECTOR

Let a cartesian coordinate system be defined in the customary manner for each bond of the chain backbone, its axes being fixed in relation to the bond in question and the preceding bond (3). The axis $x_i$ of the coordinate system for bond $i$ is assigned the direction of the $i$-th bond; the $y_i$ axis is in the plane defined by bonds $i$–1 and $i$, its direction being chosen to make an acute angle with bond $i$–1; and axis $z_i$, perpendicular to the plane of the bond pair, is assigned the direction that completes a right-handed coordinate system. In order to define a coordinate system for the first bond of the chain, let a hypothetical zeroth bond be introduced to provide a predecessor in the chain. This bond is placed in the plane of bonds 1 and 2 with its direction parallel to the latter, i.e., the zeroth bond is *trans* to the second bond, as shown in Fig. 1. Coordinate axes $y_1$ and $z_1$ may then be defined according to the foregoing conventions. It will be observed that this first coordinate system is in fact fixed with reference to bonds 1 and 2, the zeroth bond having been added merely for the purpose of defining this coordinate system in a manner equivalent to the definition of all of the succeeding ones.

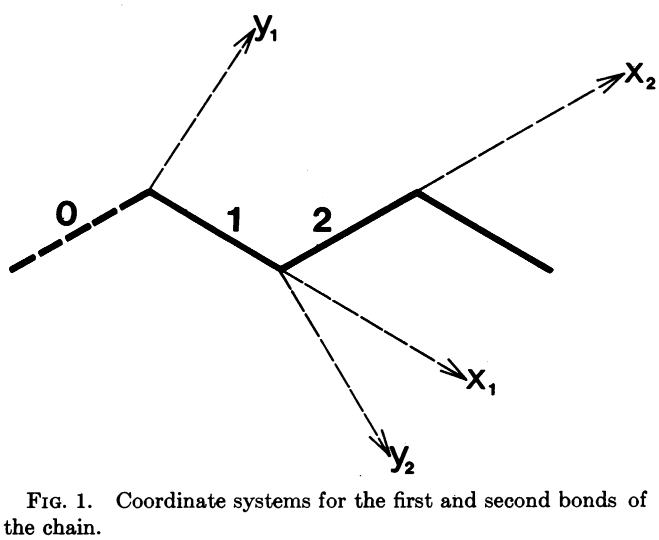

FIG. 1. Coordinate systems for the first and second bonds of the chain.

Let $x$, $y$, and $z$ be the components of the chain vector $\mathbf{r}$ expressed in the reference frame for the first bond. In column form

$$\mathbf{r} = \begin{bmatrix} x \\ y \\ z \end{bmatrix} \equiv \text{col}\,(x, y, z) \qquad [1]$$

The vector $\mathbf{r}$ is, of course, the sum of the individual bond vectors $\boldsymbol{\ell}_i$. Expressed in their respective reference frames, each takes the form $\boldsymbol{\ell}_i = \text{col}\,(l_i, 0, 0)$. For formation of the sum it is required to transform each bond vector to the first reference frame. Let $\mathbf{T}_i$ be the operator that, by premultiplication, transforms a vector in reference frame $i + 1$ to its representation in reference frame $i$. We observe that $\mathbf{T}_i = \mathbf{T}_i(\theta_i, \varphi_i)$, where $\theta_i$ is the angle between bond vectors $i$ and $i + 1$ and $\varphi_i$ is the angle of torsion about bond $i$. If all $\theta_i$ and $\varphi_i$ are specified, then $\mathbf{r}$ is determined. The required sum for a chain of $n$ bonds may then be obtained, by adaptation of methods elaborated previously (3, 6–8), as the serial product

$$\mathbf{r} = \prod_{i=1}^{n} \mathbf{A}_i \qquad [2]$$

where the $\mathbf{A}_i$ are matrices defined by

$$\mathbf{A}_i = \begin{bmatrix} \mathbf{T} & \boldsymbol{\ell} \\ 0 & 1 \end{bmatrix}_i, 1 < i < n \qquad [3]$$

$$\mathbf{A}_1 = [\mathbf{T}\ \boldsymbol{\ell}]_1 \qquad [3']$$

$$\mathbf{A}_n = \begin{bmatrix} \boldsymbol{\ell} \\ 1 \end{bmatrix}_n \qquad [3'']$$

Serial subscripts appended to the supermatrices apply to quantities $\mathbf{T}$ and $\boldsymbol{\ell}$ within the brackets. It will be understood that the $\mathbf{T}_i$ are expressed as $3 \times 3$ matrices and the bond vectors $\boldsymbol{\ell}_i$ as columns. Using a previous convention (3), we express a serial product of factors such as the $\mathbf{A}_i$ by appending a subscript to denote the index of the first factor and a superscript in parentheses to denote the number of factors. Then

$$\mathbf{r} = \mathbf{A}_1^{(n)} \qquad [4]$$

replaces Eq. [**2**].

Evaluation of the configurational average of $\mathbf{r}$ expressed in the coordinate system affixed in the molecule can be carried out by application of general methods described in detail elsewhere (3, 7, 8). Dismissing the trivial case in which the conformations of neighboring bonds are mutually independent, we take for granted that the configuration integral over various conformations may be approximated by summation over a judiciously chosen set of rotational isomeric states. The matrix of statistical weights relating the $\nu_i$ states designated for bond $i$ to those of the preceding bond, or of a sequence of several preceding bonds if neighbor interactions extend over a greater range, is represented by $\mathbf{U}_i$. If the range of interdependence does not extent beyond immediate neighbors, a condition usually met in good approximation, then the order of $\mathbf{U}_i$ is $\nu_{i-1} \times \nu_i$ for $1 < i < n$. In most instances, all bonds of the chain are of the same character or, at any rate, each is appropriately represented by the same number $\nu$ of rotational states. Then $\mathbf{U}_i$ is of square order ($i = 1$ and $n$ excepted). The configuration partition function $Z$ is given in any case by the serial product of statistical weight matrices i.e., (8)

$$Z = \mathbf{U}_1^{(n)} \qquad [5]$$

where

$$\left.\begin{aligned} \mathbf{U}_1 &= \text{row}\,(1, 0, \ldots 0) \\ \mathbf{U}_n &= \text{col}\,(1, 1, \ldots 1) \end{aligned}\right\} \qquad [6]$$

By application of general methods published previously (3, 6–8),

$$\langle \mathbf{r} \rangle = Z^{-1}\mathcal{A}_1^{(n)} \qquad [7]$$

where angle brackets denote configurational averages, the generator matrix for an internal bond is defined by (7, 8)

$$\mathcal{A}_i = (\mathbf{U}_i \otimes E_4)\|\mathbf{A}_i\|, \quad 1 < i < n \qquad [8]$$

where $\otimes$ signifies the matrix direct product, $\mathbf{E}_4$ is the identity of the same order (four) as $\mathbf{A}$, and $\|\mathbf{A}_i\|$ is the diagonal array of the $\nu_i$ matrices $\mathbf{A}_i$ for the respective rotational states of bond $i$. For the terminal bonds

$$\mathcal{A}_1 = \mathbf{U}_1 \otimes \mathbf{A}_1 \qquad [8']$$

$$\mathcal{A}_n = \mathbf{U}_n \otimes \mathbf{A}_n \qquad [8'']$$

The configurational average of the chain vector in the initial reference frame of the chain is a quantity of foremost importance. We call it the *persistence vector* and denote it by $\mathbf{a}$. That is

$$\mathbf{a} \equiv \langle \mathbf{r} \rangle = \begin{bmatrix} \langle x \rangle \\ \langle y \rangle \\ \langle z \rangle \end{bmatrix} \qquad [9]$$

For a symmetric chain, i.e., one devoid of centers of asymmetry, $\langle z \rangle = 0$. In general, $\langle y \rangle$, as well as $\langle x \rangle$, is nonzero. Their magnitudes may be expected to depend on the structure and configurational statistics of the given chain in ways that are difficult to envisage in advance of numerical calculations.*

The persistence vector may be regarded as the elaboration of the persistence length $a$ of the Porod–Kratky model chain (5). Differences should be noted carefully, however. The one is a vector whose direction is dictated by structure and spatial configuration; the other is the scalar persistence in the direc-

---

* The vector $\mathbf{r}_{hk}$ connecting atoms $h$ and $k$ located within the chain ($0 < h < k < n$) may be treated similarly. Thus, in analogy to Eq. [**2**],

$$\mathbf{r}_{hk} = {}^{\dagger}\mathbf{A}_{h+1}\mathbf{A}_{h+2}^{(k-h-2)}\mathbf{A}_k{}^{\dagger}$$

where the initial factor ${}^{\dagger}\mathbf{A}_{h+1}$ for the bond following atom $h$ and the final factor $\mathbf{A}_k{}^{\dagger}$ for the bond preceding atom $k$ are expressly noted by superscripted daggers$^{\dagger}$ preceding and following the symbols for the matrices which are defined according to Eqs. [**3′**] and [**3″**], respectively. Owing to end effects, moments of this vector may differ from those of the vector $\mathbf{r}$ connecting the ends of a chain of the same length $n = k - h$. Such effects usually are small. They are taken into account through replacement of Eq. [**7**] by

$$\langle \mathbf{r}_{hk} \rangle = Z^{-1}(\mathbf{U}_1{}^{(h)} \otimes \mathbf{E}_3)\ {}^{\dagger}\mathcal{A}_{h+1}\mathcal{A}_{h+2}^{(k-h-2)}\mathcal{A}_h{}^{\dagger}\ \mathbf{U}_{k+1}^{(n-k)} \qquad [7^*]$$

where the $\mathcal{A}_{h+2}$, etc., are defined according to Eq. [**8**] and

$${}^{\dagger}\mathcal{A}_{h+1} = [(\mathbf{U} \otimes \mathbf{E}_3)\ \|{}^{\dagger}\mathbf{A}\|]_{h+1}$$

$$\mathcal{A}_k{}^{\dagger} = \mathbf{U}_k \otimes \mathbf{A}_k{}^{\dagger}$$

For $h = 0$ and $k = n$, Eq. [**7***] reduces to Eq. [**7**]. Second and higher moments of $\mathbf{r}_{hk}$ may be treated in the manner of Eqs. [**10–15**].

tion of the initial bond, or element of arc, of a hypothetical model. In the case of real chains, $a$ has been identified (3) with the average persistence in the direction of the first bond of the chain. In this sense, $a$ is a particular component of $\mathbf{a}$.

Whereas the persistence length $a$ has been defined as the persistence in the limit of an infinitely long chain, we define $\mathbf{a}$ as the persistence vector for a chain of specified length $n$. Thus, $\mathbf{a}$ is a function of $n$. Reasons for this definition will be apparent below. Of course, for any chain of finite flexibility, $\mathbf{a}$ reaches a limit with increase in $n$.

## SECOND MOMENTS OF r

Consider the dyadic product $\mathbf{r}\mathbf{r}^T$ whose elements comprise all quadratic terms formed from the components of $\mathbf{r}$. The elements of this product form a symmetric matrix. Arranged in column form, they are given identically by $\mathbf{r} \otimes \mathbf{r}$. If identical elements are combined, this operation being denoted by an asterisk, then

$$(\mathbf{r} \otimes \mathbf{r})^* = \mathrm{col}\ (x^2, 2xy, 2xz, y^2, 2yz, z^2) \qquad [10]$$

It follows from Eq. [4] and the theorem on direct products that

$$\mathbf{r} \otimes \mathbf{r} = \mathbf{A}_1{}^{(n)} \otimes \mathbf{A}_1{}^{(n)} = (\mathbf{A} \otimes \mathbf{A})_1{}^{(n)} \qquad [11]$$

or, if the self-direct products $\mathbf{A}_i \otimes \mathbf{A}_i$ are condensed according to established methods (7)

$$(\mathbf{r} \otimes \mathbf{r})^* = [(\mathbf{A} \otimes \mathbf{A})^*]_1{}^{(n)} \qquad [12]$$

Application of averaging procedures (7, 8) yields

$$\langle(\mathbf{r} \otimes \mathbf{r})^*\rangle = Z^{-1}(\mathcal{A}^*{}_{(2)})_1{}^{(n)} \qquad [13]$$

where

$$(\mathcal{A}^*{}_{(2)})_i = (\mathbf{U} \otimes \mathbf{E}_{10})_i \| (\mathbf{A} \otimes \mathbf{A})_i{}^* \|,\ 1 < i < n \qquad [14]$$

$$(\mathcal{A}^*{}_{(2)})_1 = \mathbf{U}_1 \otimes (\mathbf{A}_1 \otimes \mathbf{A}_1)^* \qquad [14']$$

$$(\mathcal{A}^*{}_{(2)})_n = \mathbf{U}_n \otimes (\mathbf{A}_n \otimes \mathbf{A}_n)^* \qquad [14'']$$

## MOMENTS OF r OF HIGHER ORDERS

The foregoing results admit of immediate generalization to moments of any order $p$. Thus

$$\langle(\mathbf{r}^{\times p})^*\rangle = Z^{-1}(\mathcal{A}^*{}_{(p)})_1{}^{(n)}$$

where the self-direct product of degree $p$ is denoted by the superscript$^{\times p}$, and

$$(\mathcal{A}^*{}_{(p)})_i = (\mathbf{U} \otimes \mathbf{E}_\gamma) \| (\mathbf{A}_i{}^{\times p})^* \|,\ 1 < i < n \qquad [15]$$

Terminal generator matrices are defined in analogy to Eqs. [14′] and [14″].

The order of the identity to $E_\gamma$ is

$$\gamma = \binom{p+3}{p} \qquad [16]$$

and the order of $(\mathcal{A}_p{}^*)_i$ is $\nu\gamma$. If, for example, each bond is described by $\nu = 3$ rotational states and $p = 4$, then $3\gamma = 105$. If the chain is symmetric, the order may be reduced by a factor somewhat smaller than $^2/_3$ (7).

## THE VECTOR ϱ AND ITS MOMENTS

Obviously the density distribution function $W(\mathbf{r})$ is not symmetric about $\mathbf{r} = 0$ when $\mathbf{r}$ is expressed in the reference frame attached to the molecule. Only if $\mathbf{r}$ were expressed in an external coordinate system and all orientations of the molecule were averaged out would $W(\mathbf{r})$ be spherically symmetric; even so, the resulting distribution would be acentric if $|\mathbf{a}|$ is appreciable compared to $\langle r^2\rangle^{1/2}$. Representation of $W(\mathbf{r})$ by a Gaussian function centered at $\mathbf{r} = 0$ therefore entails error. This error should be most evident for short chains and for chains of intermediate length; it must vanish in importance as $n \to \infty$.

These circumstances recommend consideration of the distribution of the remote end of the chain relative to the terminus of the persistence vector. We therefore introduce the vector $\boldsymbol{\varrho}$ defined by

$$\boldsymbol{\varrho} = \mathbf{r} - \mathbf{a} \qquad [17]$$

Let its elements in the coordinate system of the first bond be $u$, $v$, $w$: i.e.,

$$\boldsymbol{\varrho} = \begin{bmatrix} u \\ v \\ w \end{bmatrix} = \begin{bmatrix} x - \langle x\rangle \\ y - \langle y\rangle \\ z - \langle z\rangle \end{bmatrix} \qquad [18]$$

By definition

$$\langle\boldsymbol{\varrho}\rangle = \mathbf{0} \qquad [19]$$

The average of its squared magnitude is

$$\langle\rho^2\rangle = \langle\boldsymbol{\varrho}^T\boldsymbol{\varrho}\rangle = \langle r^2\rangle - a^2 \qquad [20]$$

The average of the dyadic product, i.e.,

$$\langle\boldsymbol{\varrho}\boldsymbol{\varrho}^T\rangle = \langle(\mathbf{r} - \mathbf{a})(\mathbf{r} - \mathbf{a})^T\rangle \qquad [21]$$

comprises the moments of second order formed from the components of $\boldsymbol{\varrho}$. These moments are related to the corresponding moments of $\mathbf{r}$ according to the familiar relationships

$$\left.\begin{array}{l} \langle u^2\rangle = \langle x^2\rangle - \langle x\rangle^2 \\ \langle uv\rangle = \langle xy\rangle - \langle x\rangle\langle y\rangle \\ \text{etc.} \end{array}\right\} \qquad [22]$$

Higher moments of the components of $\boldsymbol{\varrho}$ are also related to moments of the components of $\mathbf{r}$, and may be obtained from the latter; e.g.,

$$\begin{array}{l} \langle u^3\rangle = \langle x^3\rangle - 3\langle x^2\rangle\langle x\rangle + 2\langle x\rangle^3 \\ \text{etc.} \end{array} \qquad [23]$$

Thus, computation of the averaged moments of the components of $\mathbf{r}$ furnishes the corresponding moments of $\boldsymbol{\varrho}$.

For chains of very great length, the higher moments dominate the lower and the need to distinguish second and higher moments of $\boldsymbol{\varrho}$ from those of $\mathbf{r}$ vanishes.

Since the matrix $\langle\boldsymbol{\varrho}\boldsymbol{\varrho}^T\rangle$ of the moments of second degree is symmetric,

$$\mathbf{R}\langle\boldsymbol{\varrho}\boldsymbol{\varrho}^T\rangle\mathbf{R}^T = \mathrm{diag}\ (\langle\rho_1{}^2\rangle, \langle\rho_2{}^2\rangle, \langle\rho_3{}^2\rangle) \qquad [24]$$

where $\mathbf{R}$ is the appropriate orthogonal transformation. The subscripts 1, 2, 3 will be used henceforth to denote the principal axes of the matrix of second moments of $\boldsymbol{\varrho}$. Components of vectors, e.g., $\mathbf{r}$, $\mathbf{a}$, and $\boldsymbol{\varrho}$, referred to these axes will be subscripted by the respective numerals.

We note in passing that the moments of second order are given in column form by the self-direct product of $\boldsymbol{\varrho}$, i.e.,

$$\langle\boldsymbol{\varrho}\boldsymbol{\varrho}^T\rangle^C = \boldsymbol{\varrho} \otimes \boldsymbol{\varrho} \equiv \boldsymbol{\varrho}^{\times 2} \qquad [25]$$

which, of course, can be condensed to $6 \times 1$ order by combining redundant elements as pointed out above.

## THE DENSITY DISTRIBUTION OF $\boldsymbol{\rho}$

In Gaussian approximation the density distribution of $\boldsymbol{\rho}$ is determined by the matrix $\langle \boldsymbol{\rho}\boldsymbol{\rho}^T \rangle$; i.e.,

$$W_a(\boldsymbol{\rho}) = (\det \boldsymbol{\beta})^{1/2}\, \pi^{-3/2} \exp(-\, \boldsymbol{\rho}^T \boldsymbol{\beta} \boldsymbol{\rho}) \qquad [26]$$

where

$$\boldsymbol{\beta} = \frac{1}{2} \langle \boldsymbol{\rho}\boldsymbol{\rho}^T \rangle^{-1} \qquad [27]$$

The subscript $a$ distinguishes the function $W_a$ representing the density distribution about the terminus of $\mathbf{a}$ from the distribution $W$ of $\mathbf{r}$ about the beginning of the sequence of units comprising the chain. Thus,

$$W_a(\boldsymbol{\rho}) \equiv W_a(\mathbf{r} - \mathbf{a}) = W(\mathbf{r}).$$

Expressed as in Eq. [26], the distribution function $W_a$ is invariant to axis rotations. One may regard it as the product of three one-dimensional Gaussians, $w_{a,1}(\rho_1)$, $w_{a,2}(\rho_2)$, $w_{a,3}(\rho_3)$, for each of the principal axes of the matrix $\boldsymbol{\beta}$ of the quadratic form in the exponent in Eq. [26]. These component distributions are characterized by the variances $\langle \rho_1{}^2 \rangle$, $\langle \rho_2{}^2 \rangle$, and $\langle \rho_3{}^2 \rangle$, respectively.

For the purpose of assessing departures from the Gaussian distribution on a uniform basis, we introduce the reduced vector $\tilde{\boldsymbol{\rho}}$ defined by

$$\tilde{\boldsymbol{\rho}} = \langle \boldsymbol{\rho}\boldsymbol{\rho}^T \rangle^{-1/2} \boldsymbol{\rho} \qquad [28]$$

Its components in the principal axes of the second moments of $\boldsymbol{\rho}$ are $\rho_1/\langle \rho_1{}^2 \rangle^{1/2}$, etc. It follows that

$$\langle \tilde{\boldsymbol{\rho}}\tilde{\boldsymbol{\rho}}^T \rangle = \mathbf{E} \qquad [29]$$

or, in column form

$$\langle \tilde{\boldsymbol{\rho}}^{\times 2} \rangle = \mathbf{E}^C \qquad [29']$$

where $\mathbf{E}$ is the identity of third order and $\mathbf{E}^C = \text{col}\ (1\ 0\ 0\ 0\ 1\ 0\ 0\ 0\ 1)$. These results are, of course, invariant to axis rotation. This spherical symmetry will not hold for the higher moments of $\tilde{\boldsymbol{\rho}}$ in general.

Substitution of Eq. [28] in [26] yields

$$W_a(\tilde{\boldsymbol{\rho}}) = (2\pi)^{-3/2} \exp\left(-{}^1\!/_2\, \tilde{\boldsymbol{\rho}}^T \tilde{\boldsymbol{\rho}}\right) \qquad [30]$$

where $W_a(\tilde{\boldsymbol{\rho}})$ is expressed per unit range in the coordinates of $\tilde{\boldsymbol{\rho}}$.

The departures that may occur relative to the simple Gaussian representation of $W_a(\boldsymbol{\rho})$ according to Eq. [26], or of $W_a(\tilde{\boldsymbol{\rho}})$ according to Eq. [30], obviously are manifold. Hence, they elude succinct characterization in general. For definiteness, therefore, consider the distribution $w_h(h)$ with respect to the single coordinate $h$ along an arbitrary axis defined by the unit vector $\mathbf{n}$. Thus, let

$$h = \mathbf{n}^T \tilde{\boldsymbol{\rho}} \qquad [31]$$

The distribution $w_h(h)$ of $h$ is the integral of the density distribution $W_a(\tilde{\boldsymbol{\rho}})$ of $\tilde{\boldsymbol{\rho}}$ over the full ranges of the coordinates transverse to $\mathbf{n}$. Moments of $h$ are given by

$$\langle h^p \rangle = \langle (\mathbf{n}^T \tilde{\boldsymbol{\rho}})^p \rangle = (\mathbf{n}^{\times p})^T \langle \tilde{\boldsymbol{\rho}}^{\times p} \rangle \qquad [32]$$

where $\langle \tilde{\boldsymbol{\rho}}^{\times p} \rangle = (\langle \boldsymbol{\rho}\boldsymbol{\rho}^T \rangle^{-1/2})^{\times p} \langle \boldsymbol{\rho}^{\times p} \rangle$. According to the definitions of $\boldsymbol{\rho}$ and $\tilde{\boldsymbol{\rho}}$, $\langle h \rangle = 0$ and $\langle h^2 \rangle = 1$.

Let $\delta_1, \delta_2, \ldots \delta_p$, etc., be defined as follows:

$$\left.\begin{aligned} \delta_1 &= \delta_2 = 0 \\ \delta_3 &= \langle h^3 \rangle = \langle (\mathbf{n}^T \tilde{\boldsymbol{\rho}})^3 \rangle = (\mathbf{n}^{\times 3})^T \langle \tilde{\boldsymbol{\rho}}^{\times 3} \rangle \\ \delta_4 &= \langle h^4 \rangle - 3 \\ \delta_5 &= \langle h^5 \rangle \\ \delta_6 &= \langle h^6 \rangle - 3 \cdot 5 \\ &\text{etc.} \end{aligned}\right\} \qquad [33]$$

If $W_a(\boldsymbol{\rho})$ is Gaussian, $\delta_p = 0$ for all $p$. Hence, the distribution $w_h(h)$ referred to any axis $\mathbf{n}$ through the terminus of $\mathbf{a}$ is Gaussian.

To complete the analysis we expand $w_h(h)$ in the Hermite polynomials $H_0(q) = 1$, $H_1(q) = 2q$, $H_2(q) = (2q)^2 - 2$, etc. Then

$$w_h(h) = (2\pi)^{-1/2} \exp\left(-h^2/2\right) \sum_{m=0}^{\infty} C_m H_m(h/\sqrt{2}) \qquad [34]$$

The coefficients $C_m$ evaluated in terms of the moments of $h$ by resort to the familiar orthogonality condition are

$$\left.\begin{aligned} C_0 &= 1,\ C_1 = C_2 = 0 \\ C_3 &= \delta_3/2^{3/2} \cdot 3! \\ C_4 &= \delta_4/2^2 \cdot 4! \\ C_5 &= (\delta_5 - 10\delta_3)/2^{5/2} \cdot 5! \\ C_6 &= (\delta_6 - 15\delta_4)/2^3 \cdot 6! \end{aligned}\right\} \qquad [35]$$

Investigations of $w_h(h)$ in this manner with $\mathbf{n}$ taken along each of the principal axes of the second moments of $\boldsymbol{\rho}$ may be expected to lead to a reliable appraisal of the degree of departure from the Gaussian distribution expressed by Eq. [26] or [30].

## CONCLUDING REMARKS

Often the radial distribution function $R(r)$ will be required, this being the integral of $W(\mathbf{r}) = W_a(\boldsymbol{\rho})$, with $\boldsymbol{\rho} = \mathbf{r} - \mathbf{a}$, over all directions for the fixed value $r$ of the magnitude of $\mathbf{r}$. Although $R(r)$ is determined by $W_a(\boldsymbol{\rho})$, evaluation of the former from the latter may involve extensive numerical calculations. Only in the unlikely event that $W_a(\boldsymbol{\rho})$ is both symmetric about $\mathbf{a}$ and Gaussian as well, is the conversion expressible in simple mathematical form. The task may be simplified, however, by resort to certain approximations that suggest themselves as expedients. Demonstration of their efficacy must await evaluations of the moments of $\mathbf{r}$ and of $\boldsymbol{\rho}$ for representative examples of real polymer chains, followed by examination of the characteristics of the density distribution function.

The distinction between $W(\mathbf{r})$ about $\mathbf{r} = \mathbf{0}$ [and the associated radial distribution function $R(r)$] on the one hand and the distribution function $W_a(\boldsymbol{\rho})$ about $\mathbf{r} = \mathbf{a}$ on the other should be important for chains of lengths such that $\langle r^2 \rangle^{1/2}$ does not greatly exceed the magnitude of $\mathbf{a}$. The acentricity implied in this distinction may account for the discrepancies between results of Monte Carlo calculations applied to *n*-alkane chains by Fixman and Alben (9) and the customary, spherically symmetric representations (Gaussian) of the density distributions. Their Monte Carlo densities for chains having fewer than about 30 bonds fall well below those given by the latter representations. Similar results have been found by Beevers and Semlyen (10) for short poly(dimethylsiloxane) chains.

Marked departures from spherical symmetry of $W(\mathbf{r})$ are indicated for polynucleotide chains according to Monte Carlo calculations of DeLisi and Crothers (11, 12). For chains of fewer than twenty residues (120 skeletal bonds), $W(\mathbf{r})$ was found to depart drastically from a spherically symmetric Gaussian. Expansion in a Hermite series of the even moments of $\mathbf{r}$ up to $\langle r^{40} \rangle$ was insufficient to overcome the departure (12). Marked acentricity in the chain vector distribution for short polynucleotide chains is consistent with the fairly large extension that is characteristic of these chains in the random-coil state (13, 14). The matter bears directly on the analysis of single-strand loops such as occur in t-RNA, i.e., on the loop weighting functions treated by DeLisi and Crothers (11, 12).

The extent to which these discrepancies between distribution functions based on moments and those obtained from Monte Carlo calculations (or from properties of real chains) can be resolved through development of the distribution about the terminus of the persistence vector remains to be established.

This work was supported by the Directorate of Chemical Sciences, U.S. Air Force Office of Scientific Research Grant no. AFOSR-73-2441.

1. Volkenstein, M. V. (1963) *Configurational Statistics of Polymeric Chains* (translated from the Russian ed. Timasheff, S. N. & Timasheff, M. J.) (Interscience, New York).
2. Birshtein, T. M. & Ptitsyn, O. B. (1966) *Conformations of Macromolecules* (translated from the 1964 Russian ed., Timasheff, S. N. & Timasheff, M. J.) (Interscience, New York).
3. Flory, P. J. (1969) *Statistical Mechanics of Chain Molecules* (Interscience, New York).
4. Nagai, K. (1968) *J. Chem. Phys.* **49**, 4212–4215; see also (1970) *Report No.* **334** (Government Industrial Research Institute, Osaka, Japan).
5. Porod, G. (1949) *Monatsh. Chem.* **80**, 251–255; Kratky, O. & Porod, G. (1949) *Rec. Trav. Chim.* **68**, 1106–1122.
6. Flory, P. J. (1964) *Proc. Nat. Acad. Sci. USA* **51**, 1060–1067.
7. Flory, P. J. & Abe, Y. (1971) *J. Chem. Phys.* **54**, 1351–1363.
8. Flory, P. J. (1972) *J. Chem. Phys.* **56**, 862–866.
9. Fixman, M. & Alben, R. (1973) *J. Chem. Phys.* **58**, 1553.
10. Beevers, M. S. & Semlyen, J. A. (1972) *Polymer* **13**, 385–390.
11. DeLisi, C. & Crothers, D. M. (1971) *Proc. Nat. Acad. Sci. USA* **68**, 2682–2685; (1971) *Biopolymers* **10**, 1809–1827.
12. DeLisi, C. (1972) *Biopolymers* **11**, 2251–2265.
13. Eisenberg, H. & Felsenfeld, G. (1967) *J. Mol. Biol.* **30**, 17–37; Inners, L. D. & Felsenfeld, G. (1970) *J. Mol. Biol.* **50**, 373–389; Achter, E. K. & Felsenfeld, G. (1971) *Biopolymers* **10**, 1625–1634.
14. Olson, W. K. & Flory, P. J. (1972) *Biopolymers* **11**, 1–23, 25–26, 57–66.