Computational and Theoretical Polymer Science 11 (2001) 143–153

Computational
and Theoretical
Polymer Science

www.elsevier.nl/locate/ctps

Conformational statistics of macromolecules using generalized convolution

G.S. Chirikjian

Department of Mechanical Engineering, Johns Hopkins University, Baltimore, MD 21218, USA

Received 8 April 1999; received in revised form 6 October 1999; accepted 7 October 1999

Abstract

A new technique for generating statistical properties of chain-molecule conformations is presented. Conditional probability density functions (PDFs) describing the frequency of occurrence of the relative position and orientation of frames of reference affixed to selected backbone atoms serve as the inputs. Ensemble statistical properties of whole chains are generated by performing multiple generalized convolutions of these conditional PDFs. The formulation is shown to include classical theories such as the hindered and freely rotating chains, the Gaussian random walk, and the rotational isomeric state model. The convolution model is modified to include the long-range effects of excluded volume. An analytical example is used to illustrate the procedure. A general algorithm to calculate the ensemble properties of an arbitrary chain macromolecule is presented. In this algorithm, each of the N degrees of freedom (e.g. torsion angles) is assumed to have K discrete states. Using the convolution procedure, a chain is divided into P statistical units. The computational requirement is reduced from an $\mathcal{O}(K^N)$ calculation (corresponding to direct enumeration) to one which is $\mathcal{O}(P(C + K^{N/P}))$ where C is the computational complexity of the convolution procedure. In the case of a homopolymer, computations are reduced further to $\mathcal{O}(C \log(P) + K^{N/P})$. © 2000 Elsevier Science Ltd. All rights reserved.

Keywords: Macromolecule; Conformation; Ensemble; Convolution; Rigid-body motion; Probability density function

1. Introduction

The study of conformation-dependent properties of macromolecular systems is quite old. The use of statistical techniques in polymer chemistry dates back to the 1930s with the work of Kuhn [1] and was extensively explored in the 1950s [3–5]. The rotational isomeric state (RIS) model was popularized in the 1960s with the advent of Flory's classic book [6], and continues to be of interest today [9]. Work in the area of statistical analysis of macromolecular conformations has received much attention up to the current day (see, e.g. Refs. [7,15,18–22,25–29]). A number of idealized models such as Gaussian random walks, jointed chains (with or without fixed bond angles and hindered torsion angles), and worm-like chains have been explored in great detail, and are explained thoroughly in classic books in polymer science [6,8,10–14].

As a general rule, the analysis is easiest when the effects of conformational energy can be disregarded (i.e. when all conformations can be considered to have the same energy). In this case, purely kinematical models such as the freely jointed chain are used. The next level of sophistication is when the energy of local interactions is modeled. Then comes the RIS model in which pairwise interdependent potentials are used. At another level of approximation, the longer range effects of excluded volume are modeled using self-avoiding walks [20,28]. An altogether different approach is to do direct molecular dynamic simulations (see Refs. [23,24] for a recent example). The primary drawback of molecular simulations is the intensive computational requirements.

Regardless of which technique is used, a potential function that closely models the physical reality is quite important. One of the most popular choices is the Lennard-Jones potential [16]. Due to the high computational cost of direct molecular simulation, one of the most popular approaches in conformational analysis of macromolecules is the Monte Carlo method [17]. While Monte Carlo sampling is a very general and powerful technique, it too has some drawbacks, the most obvious of which is its poor performance in capturing the "tails" of certain kinds of PDFs in polymer science.

The approach taken in this paper is to apply an altogether different statistical analysis of conformational properties of *macroscopically serial*[1] chain molecules. The basic concept

---

[1] A macromolecule which may contain local loops or branches, but appears serial on large enough length scales (e.g. DNA).

PII: S1089-3156(99)00086-0

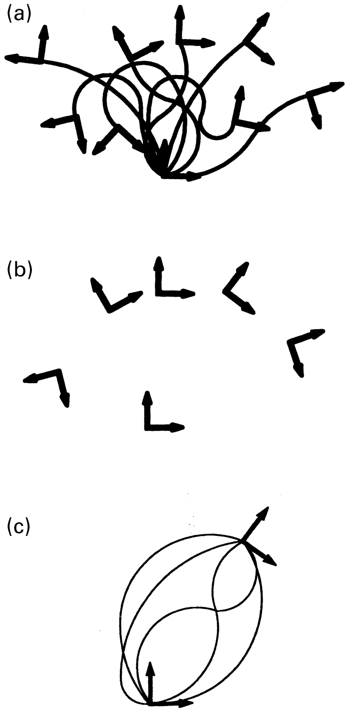

Fig. 1. The functions (a) $\rho(\bar{x})$, (b) $f(g)$ and (c) $\mu(g, \bar{x})$.

behind this work is not new: a number of reference frames are affixed to the backbone of the molecule, and the statistical properties for each segment connecting neighboring frames are independently generated. The novelty of this formulation results from the recognition that a generalized convolution (in the sense defined in the next section) of frame distributions generates the statistics for the whole molecule. This distribution contains all the information commonly calculated in polymer science, such as mean squared radius of gyration and mean squared end-to-end distance. The benefit of the method is that it also provides a tool for naturally including the effects of conformational energy and excluded volume (at least qualitatively) while retaining the simplicity of a statistical analysis.

The remainder of this paper is structured as follows: Section 2 introduces the concept of frame distributions and reviews the mathematics of generalized convolutions. Section 3 shows how this concept can be used to generate macromolecule properties when the effects of excluded volume and other long-range effects can be ignored, and illustrates the technique with analytical examples. In Section 4, long-range interactions are modeled and incorporated into the framework. In Section 5, closed-form analytical examples based on the mathematical formulation of this paper are described.

2. Mass density, frame density, and generalized convolutions

In this section we define three statistical properties of macromolecular ensembles. Later it will be shown how these quantities are used to model the interactions of segments which may be far apart as measured along the chain, but are proximal in space. Before introducing these definitions, some notation is required.

Let $\mathbb{R}^3$ denote three-dimensional space, and G denote the set of all rigid-body motions. $\bar{x} \in \mathbb{R}^3$ is a position vector, and $g = (R, \bar{r}) \in G$ is a rigid-body transformation where R is a 3×3 rotation matrix and $\bar{r}$ is a translation vector. Sometimes it is convenient to represent a rigid-body motion in three-dimensional space as a 4×4 homogeneous transformation matrix:

$$H(g) = \begin{pmatrix} R & \bar{r} \\ \bar{0}^{\mathrm{T}} & 1 \end{pmatrix}$$

(see, e.g. Ref. [9. p. 114]).

Two rigid-body motions are composed as $g_1 \circ g_2 = (R_1 R_2, R_1 \bar{r}_2 + \bar{r}_1)$, which is the same as performing the matrix product $H(g_1)H(g_2) = H(g_1 \circ g_2)$.

The composition of rigid-body motions can also be viewed geometrically as a sequential change of reference frames. The inverse of any $g \in G$ and the identity (do nothing) motion are given, respectively, as $g^{-1} = (R^{\mathrm{T}}, -R^{\mathrm{T}}\bar{r})$ and $e = (I, \bar{0})$ where I is the 3×3 identity matrix. The elements of G act on position vectors in $\mathbb{R}^3$ as $g \circ \bar{x} = R\bar{x} + \bar{r}$. A more detailed review of rigid-body motions, including invariant integration and convolution of functions of motion is presented in Appendix A.

The three statistical quantities of importance in the present formulation are: (1) the ensemble mass density for the whole chain $\rho(\bar{x})$;(2) the ensemble tip frame density $f(g)$ (where g is the frame of reference of the distal end of the chain relative to the proximal); (3) the function $\mu(g, \bar{x})$, which is the ensemble mass density of all configurations which grow from the identity frame fixed to one end of the chain and terminate at the relative frame g at the other end. These quantities can be thought of as histograms on $\mathbb{R}^3$, G, and $G \times \mathbb{R}^3$ (this last quantity is the set of all pairs of the form $(g, \bar{x})$ for $g \in G$ and $\bar{x} \in \mathbb{R}^3$) corresponding to the collections of conformations illustrated in Fig. 1.

The functions ρ, f, and μ are related to each other. Given $\mu(g, \bar{x})$, the ensemble mass density is calculated by adding the contribution of each μ for each different end position and orientation:

$$\rho(\bar{x}) = \int_G \mu(g, \bar{x}) \, \mathrm{d}g. \tag{1}$$

This integration is written as being over all motions of the end of the chain, but only frames g in the support of μ contribute to the integral. dg denotes the invariant integration measure for G reviewed in Appendix A.

In an analogous way, it is not difficult to see that integrating the $\bar{x}$-dependence out of μ provides the total mass of configurations of the chain starting at frame e (at the proximal end) and terminating at frame g (at the distal end).

Since each chain has mass M, this means that the frame density $f(g)$ is related to $\mu(g, \bar{x})$ as:

$$f(g) = \frac{1}{M} \int_{\mathbb{R}^3} \mu(g, \bar{x})\, \mathrm{d}\bar{x}. \tag{2}$$

We note the total number of frames attained by one end of the chain relative to the other is

$$F = K^N = \int_G f(g)\, \mathrm{d}g$$

when each of the chains N degrees of freedom has K preferred states. It then follows that

$$\int_{\mathbb{R}^3} \rho(\bar{x})\, \mathrm{d}\bar{x} = F{\cdot}M.$$

If the functions $\rho(\bar{x})$ and $f(g)$ are known for the whole chain then a number of important thermodynamic and mechanical properties of the polymer can be determined. For instance, the moments of any positive integer power of the end-to-end distance, $\langle |\bar{r}|^m \rangle$, can be calculated from $f(g) = f(R, \bar{r})$ by first integrating out the orientational dependence:

$$\tilde{\rho}(\bar{r}) = \int_{\mathrm{SO}(3)} f(R, \bar{r})\, \mathrm{d}R$$

where $\mathrm{d}R$ is the normalized invariant integration measure for SO(3) (the set of all 3×3 rotation matrices). Then

$$\langle |\bar{r}|^m \rangle = \int_{\mathbb{R}^3} |\bar{r}|^m \tilde{\rho}'(\bar{r})\, \mathrm{d}\bar{r}$$

where $\mathrm{d}\bar{r} = \mathrm{d}r_1\, \mathrm{d}r_2\, \mathrm{d}r_3$ is the usual integration measure for $\mathbb{R}^3$ and $\tilde{\rho}' = \tilde{\rho}/K^n$ is the normalized version of $\tilde{\rho}$. The PDF of end-to-end distances is given as

$$\mathrm{d}(r) = r^2 \int_{S^2} \tilde{\rho}'(\bar{r})\, \mathrm{d}\bar{u} \tag{3}$$

where in spherical coordinates $\bar{r} = r\bar{u}$. Here $\bar{u} \in S^2$ is a point on the unit sphere and $\mathrm{d}\bar{u}$ is the area element for the unit sphere. In the special case when $\tilde{\rho}'(\bar{r})$ is spherically symmetric $\mathrm{d}(r) = 4\pi r^2 \tilde{\rho}'(r)$. More generally, it follows that

$$\int_{\mathbb{R}^3} \tilde{\rho}'(\bar{r})\, \mathrm{d}\bar{r} = \int_0^\infty \mathrm{d}(r)\, \mathrm{d}r.$$

The functions $\rho(\bar{x})$ and $\tilde{\rho}(\bar{r})$ are also related in the following way. Imagine the chain is made up of n units,[2] each of which has K preferred energy states, and the ith of which has mass m_i. Then the function $\tilde{\rho}_i(\bar{r})$ relating the distribution of points reachable by the distal end of the ith unit relative to the proximal end of the first unit is related to the mass density of the collection of all units as

$$\rho(\bar{x}) \approx \frac{1}{K^n} \sum_{i=0}^{n-1} m_i K^{n-i} \tilde{\rho}_i(\bar{x}).$$

The weights K_{n-i} reflect the fact that units at the base of the chain will be counted more times than those at the top since they contribute to all subsequent units.

From an analytical viewpoint, all the information above can be calculated if $\mu(g, \bar{x})$ is known for the whole chain. If, however, one views the problem from a computational perspective, it is difficult to rationalize storing numerical values of the function $\mu(g, \bar{x})$ because G is a six-parameter space (three for translation and three for rotation), $\mathbb{R}^3$ is three dimensional, and the computational storage required for accurate approximation of a function on the nine-dimensional space $G \times \mathbb{R}^3$ is prohibitive. For example, a hundred sample points in each coordinate direction would require 10^{18} memory locations. However, calculating $\rho(\bar{x})$ and $f(g)$ separately with the same sampling makes the problem literally a million times easier to handle.

Moreover, there are two major problems that must be addressed in order to accurately and efficiently calculate the functions $\rho(\bar{x})$ and $f(g)$: (1) for a chain with $N \in [100, 10\,000]$ units, each neighboring pair of which has K potential wells, it is impossible with current technology to enumerate the K^N conformations of the chain corresponding to all possible combinations of potential wells; (2) the long-range interactions of units that are distant in the chain but proximal in space cannot be modeled using only the functions $\rho(\bar{x})$ and $f(g)$. Both of these problems are addressed using the recursive procedure outlined in the following sections.

3. Generating ensemble properties for purely kinematical models

This section explores mathematical and computational methods for generating statistical ensembles of macroscopically serial chain macromolecules. In particular, given ensemble properties of subchains, we explore how the properties of the whole chain can be generated if the effects of conformational energy between segments are ignored. While this can be a valid approximation in situations when the energy of interaction is considered negligible, our goal in neglecting energy effects in this section is purely a matter of pedagogy. That is, introducing useful syntax in the context of a purely kinematical model makes the definitions and concepts used in subsequent sections (where energy effects are incorporated) easier to understand.

Because of the exponential growth in the number of conformations as a function of the number of backbone atoms, it is not possible to generate the functions $\rho(\bar{x})$ and $f(g)$ for the whole chain by explicitly enumerating all configurations. This is well known in the polymer science literature. As reviewed in Section 1, the standard approaches to avoiding this exponential growth are: (1) asymptotic approximations in which the ensemble properties are assumed to be Gaussian as $N \to \infty$; and (2) Monte Carlo simulations which choose a small portion of the

---

[2] These could be either segments of the chain or single monomer units.

conformations and approximate overall behavior based on appropriate sampling.

In this section a completely different approach is taken. At the core of this approach remains the philosophy of considering ensemble properties of subsegments of the chain. This idea is not new, its roots go back more than a half century [1]. However, the way in which properties of the whole chain are calculated based on the properties of segments of the chain is quite different than traditional approaches.

The basic idea is that we imagine dividing the chain up into P statistically significant segments. P is chosen large enough so that the continuum approximations to the ensembles $\rho(\bar{x})$ and $f(g)$ meet acceptable measures of accuracy. For example, if $N \approx 1000$ and $K = 3$ (as might be the case for torsion angles in a polyethylene molecule) we might choose $P \approx 40$. In this way $K^{N/P}$ is a number which can be managed with a personal computer.

3.1. The mathematical formulation

For each of the P statistical segments in the chain we can calculate $\rho_i(\bar{x})$ and $f_i(g)$ where g is the *relative* frame of reference of the distal end of the segment with respect to the proximal one. For a homogeneous chain, such as polyethylene, these functions are the same for each value of $i = 1, \ldots, P$.

In the general case of a heterogenous chain, we can calculate the functions $\rho_{i,i+1}(\bar{x})$ and $f_{i,i+1}(g)$ for the concatenation of segments i and $i+1$ from those of segments i and $i+1$ separately in the following way:

$$\rho_{i,i+1}(\bar{x}) = F_{i+1}\rho_i(\bar{x}) + \int_G f_i(h)\rho_{i+1}(h^{-1}\circ\bar{x})\,\mathrm{d}h \tag{4}$$

and

$$f_{i,i+1}(g) = (f_i * f_{i+1})(g) = \int_G f_i(h)f_{i+1}(h^{-1}\circ g)\,\mathrm{d}h. \tag{5}$$

In these expressions $h \in G$ is a dummy variable of integration. The meaning of Eq. (4) is that the mass density of the ensemble of all conformations of the two concatenated chains segments results from two contributions. The first is the mass density of all the conformations of the lower segment (weighted by the number of different upper segments it can carry, which is $F_{i+1} = \int_G f_{i+1}\,\mathrm{d}g$). The second contribution results from rotating and translating the mass density of the ensemble of the upper segment, and adding the contribution at each of these poses (positions and orientations). This contribution is weighted by the number of frames that the distal end of the lower segment can attain relative to its base. Mathematically $L(h)\rho_{i+1}(\bar{x}) = \rho_{i+1}(h^{-1}\circ\bar{x})$ is a left-shift operation which geometrically has the significance of rigidly translating and rotating the function $\rho_{i+1}(\bar{x})$ by the transformation h. The weight $f_i(h)\,\mathrm{d}h$ is the number of configurations of the ith segment terminating at frame of reference h.

The meaning of Eq. (5) is that the distribution of frames of reference at the terminal end of the concatenation of segments i and $i+1$ is the *generalized convolution* of the frame densities of the terminal ends of each of the two segments relative to their respective bases. A more detailed explanation of why the convolution of frame densities generates the density of the concatenation of segments can be found in the author's previous work in the context of highly articulated robot arms [31]. A short review of integration over rigid-body motions and the concept of generalized convolution is presented in Appendix A. A review of the Fourier transform of functions on G (which is a tool for performing such convolutions) can be found in Ref. [30].

Eqs. (4) and (5) can be iterated with $\rho_{i,i+1}$ and $f_{i,i+1}$ taking the places of ρ_i and f_i, and $\rho_{i,i+2}$ and $f_{i,i+2}$ taking the places of $\rho_{i,i+1}$ and $f_{i,i+1}$. This is described in detail later.

Later in the paper the effects of interaction between distal segments of a macromolecule are approximated in an average sense[3] by considering how the functions $\mu_{i,i+j}(g,\bar{x})$ and $\mu_{i+j+1,k}(g,\bar{x})$ overlap for $k \geq i+j+1$. In order to calculate the interaction of these functions, one must first calculate the functions $\mu_{i,i+j}(g,\bar{x})$ from the set of functions $(\mu_{i,i+1}(g,\bar{x}), \ldots, \mu_{i+j-1,i+j}(g,\bar{x}))$. In analogy with the way $\rho_{i,i+j}(\bar{x})$ and $f_{i,i+j}(g)$ are calculated by repeating the computations in Eqs. (4) and (5), $\mu_{i,i+j}(g,\bar{x})$ is constructed recursively by first calculating $\mu_{i,i+1}(g,\bar{x})$ from $\mu_i(g,\bar{x})$ and $\mu_{i+1}(g,\bar{x})$.

This is achieved by observing that

$$\mu_{i,i+1}(g,\bar{x}) = \int_G (\mu_i(h,\bar{x})f_{i+1}(h^{-1}\circ g) + f_i(h)\mu_{i+1}(h^{-1}\circ g, h^{-1}\circ\bar{x}))\,\mathrm{d}h. \tag{6}$$

This equation says that there are two contributions to $\mu_{i,i+1}(g,\bar{x})$. The first comes from adding up all the contributions due to each $\mu_i(h,\bar{x})$. This is weighted by the number of upper segment conformations with distal ends that reach the frame g given that their base is at frame h. The second comes from adding up all shifted (translated and rotated) copies of $\mu_{i+1}(g,\bar{x})$, where the shifting is performed by the lower distribution, and the sum is weighted by the number of distinct configurations of the lower segment that terminate at h. This number is $f_1(h)\,\mathrm{d}h$.

The veracity of this derivation may be confirmed by integrating the resulting function $\mu_{i,i+1}(g,\bar{x})$ over $\mathbb{R}^N$ and G and comparing with Eqs. (4) and (5). Note that segment i has mass M_i, segment $i+1$ has mass M_{i+1}, and so $M_{i,i+1} = M_i + M_{i+1}$ is the value of M in Eq. (2) in the context of the present discussion.

3.2. Computationally efficient strategies for calculating f(g) and ρ($\bar{x}$)

Given the above means for calculating the functions

---

[3] The meaning of this "average sense" is made precise later in the paper.

$\rho_{i,i+1}(\bar{x})$ and $f_{i,i+1}(g)$ from the functions $f_i(g)$, $f_{i+1}(g)$, $\rho_i(\bar{x})$, and $\rho_{i+1}(\bar{x})$, it is now possible to formulate algorithms for generating these functions for the whole chain: $\rho(\bar{x}) = \rho_{1,P}(\bar{x})$ and $f(g) = f_{1,P}(g)$.

On a serial processor the most straightforward way to do this is to sequentially start at one end of the chain and repeatedly perform the required integrations. It does not matter at which end we begin. Starting at the base and working toward the distal end, we would calculate the sequence of functions $(\rho_{1,2}(\bar{x}), f_{1,2}(g)), \ldots, (\rho_{1,i}(\bar{x}), f_{1,i}(g)), \ldots, (\rho_{1,P}(\bar{x}), f_{1,P}(g))$. Starting from the other end we would calculate $(\rho_{P-1,P}(\bar{x}), f_{P-1,P}(g)), \ldots, (\rho_{P-i,P}(\bar{x}), f_{P-i,P}(g)), \ldots, (\rho_{1,P}(\bar{x}), f_{1,P}(g))$. In either case, viewing the number of computations required to calculate Eqs. (4) and (5) as a constant, the recursive computation of these convolution-like integrals requires $\mathcal{O}(P)$ calculations. This is after each of the functions $\rho_i(\bar{x})$ and $f_i(g)$ have been calculated, which can be achieved in $\mathcal{O}(P \cdot K^{(N/P)})$ calculations. That is, $\mathcal{O}(K^{(N/P)})$ calculations to explicitly enumerate configurations of each of the P segments.

The computational speed of the above approach on a parallel computer with P processors is much faster than on a single processor. Clearly, in this case the enumeration of segment conformations is reduced to an $\mathcal{O}(K^{(N/P)})$ time calculation since each of the P ensembles can be calculated separately. Furthermore, instead of explicitly computing a P-fold convolution requiring $\mathcal{O}(P)$ time, convolutions of adjacent functions can be calculated in a pairwise fashion on different processors. This reduces the running time to $\mathcal{O}(\log_2 P)$. For example, if $P = 8$, then the convolutions $f_{1,2} = f_1 * f_2$, $f_{3,4} = f_3 * f_4$, $f_{5,6} = f_5 * f_6$, and $f_{7,8} = f_7 * f_8$ are all performed at the same time using $4 = P/2$ processors. Then $f_{1,4} = f_{1,2} * f_{3,4}$ and $f_{5,8} = f_{5,6} * f_{7,8}$ are calculated at the next level using two processors. Finally, $f_{1,8} = f_{1,4} * f_{5,8}$ is performed using a single processor. Thus, we have in this example an eight-fold convolution calculated in the same time as $3 = \log_2(8)$ convolutions on a serial processor.

It is worth noting that the same speed-up achieved for a heterogeneous chain calculated on a parallel processor is valid in the case of a homogeneous chain calculated on a single processor. In this special case $\mathcal{O}(K^{(N/P)})$ time is required for brute force enumeration of one segments of length N/P. Similarly, a P-fold convolution of the same function with itself only requires $\mathcal{O}(\log_2 P)$ distinct convolutions, and thus this order of time. For example, the three convolutions $f_1 = f * f$, $f_2 = f_1 * f_1$, and $f_3 = f_2 * f_2$ generate the same result as an eight-fold convolution of f with itself.

4. Incorporating conformational energy effects

Including the effects of conformational energy, the density function describing the distribution of tip-to-base positions and orientations of a macromolecule may be written in the form

$$f(g)\mathrm{Vol}(\Delta(g)) = \int_{\bar{\phi} \in \mathrm{Im}(\Delta(g))} \mathrm{e}^{-E(\bar{\phi})/\mathrm{k_B}T} \, \mathrm{d}\bar{\phi} \tag{7}$$

where in the case of a serial chain $\bar{\phi} = (\phi_1, \phi_2, \ldots, \phi_{n-1})$ is the set of all torsion angles, and $\mathrm{d}\bar{\phi} = \mathrm{d}\phi_1 \ldots \mathrm{d}\phi_{n-1}$.

$\Delta(g)$ is a small six-dimensional voxel (box) in G containing g, and $\mathrm{Vol}(\Delta(g))$ is the volume of this voxel. Since the support of f is finite, it can be divided into a finite number of voxels, $\mathscr{M}$. $g(\bar{\phi})$ is the end frame of reference of the chain relative to its base for given torsion angles $\bar{\phi}$, and $\mathrm{Im}(\Delta(g))$ is the set of all torsion angles such that $g(\bar{\phi}) \in \Delta(g)$.

Each torsion angle takes its values from the unit circle, T^1, and so the whole collection of angles takes its values from the $n-1$-dimensional torus $T^{n-1} = T^1 \times \cdots \times T^1$.

One can normalize $f(g)$ in Eq. (7) by observing that the sum

$$\sum_{i=1}^{\mathscr{M}} f(g_i)\mathrm{Vol}(\Delta(g_i)) = \sum_{i=1}^{\mathscr{M}} \int_{\bar{\phi} \in \mathrm{Im}(\Delta(g_i))} \mathrm{e}^{-E(\bar{\phi})\mathrm{k_B}T} \, \mathrm{d}\bar{\phi}$$

becomes

$$\int_G f(g) \, \mathrm{d}g = \int_{\bar{\phi} \in T^{n-1}} \mathrm{e}^{-E(\bar{\phi})/\mathrm{k_B}T} \, \mathrm{d}\bar{\phi}$$

for sufficiently small voxels.

We note that as $\mathrm{k_B}T \to \infty$, $f(g)$ reduces to the purely kinematic model discussed earlier. In the following subsections we examine how the density function in Eq. (7) is related to the generalized convolution model under a wide variety of conditions.

4.1. Nearest-neighbor energy functions

One of the simplest kinds of conformational energy functions is one of the form

$$E(\bar{\phi}) = \sum_{i=1}^{n-1} E_i(\phi_i). \tag{8}$$

This kind of conformational energy function models nearest-neighbor interactions and leads to a separable partition function.

The frame density function for the concatenation of two chain molecules with this kind of energy function is again given by a generalized convolution. That is, since the energy function is additive, and the partition function is separable, the PDFs of two concatenated segments are multiplied and integrated as

$$(f_1 * f_2)(g) = f(g).$$

This is derived by substituting the energy function in Eq. (8) into Eq. (7) for a chain with $n = n_1 + n_2$ torsion angles. The functions ρ and μ are calculated analogously.

4.2. Interdependent potential functions

The rotational isomeric state (RIS) model [2] is perhaps the most widely known method to generate the statistical information needed to weight the relative occurrence of polymer conformations in a statistical mechanical ensemble. There are two basic assumptions to the RIS model developed by Flory:

- The conformational energy function for a chain molecule is dominated by interactions between each set of three groups of atoms at the intersection and ends of two adjacent bond vectors. Hence the conformational energy function can be written in the form
$$E(\bar{\phi}) = \sum_{i=2}^{n-1} E_i(\phi_{i-1}, \phi_i) \tag{9}$$
where $\bar{\phi} = (\phi_1, \phi_2, \ldots, \phi_{n-1})$ is the set of torsion angles.
- The value of the conformational partition function $\mathrm{e}^{-E(\bar{\phi})/k_\mathrm{B}T}$ is negligible except at the finite number of points where $E(\bar{\phi})$ is minimized, and hence averages of conformation-dependent functions may be calculated in the following way:
$$\langle f \rangle = \frac{\int_{T^{n-1}} f(\bar{\phi})\, \mathrm{e}^{-E(\bar{\phi})/k_\mathrm{B}T}\, \mathrm{d}\phi_1 \ldots \mathrm{d}\phi_{n-1}}{\int_{T^{n-1}} \mathrm{e}^{-E(\bar{\phi})/k_\mathrm{B}T}\, \mathrm{d}\phi_1 \ldots \mathrm{d}\phi_{n-1}} \approx \frac{\sum_{\bar{\phi}_\eta \in T^{n-1}} f(\bar{\phi}_\eta)\, \mathrm{e}^{-E(\bar{\phi}_\eta)/k_\mathrm{B}T}}{\sum_{\bar{\phi}_\eta \in T^{n-1}} \mathrm{e}^{-E(\bar{\phi}_\eta)/k_\mathrm{B}T}}.$$
Here the finite set of conformations $\{\bar{\phi}_\eta\}$ are local minima of the energy function $E(\bar{\phi})$.

As a consequence of the above two assumptions, conformational statistics are generated in the context of the RIS model using statistical weight matrices.

One drawback of this technique, as pointed out in Ref. [9], is that the actual distribution of end-to-end distance is not calculated using the RIS method. The traditional technique for generating statistical distributions is to use Monte Carlo simulations. While this is a quite effective technique, it has the drawback that it will, by definition, not pick up the tails of a rapidly decreasing distribution. In contrast, the generalized convolution technique can be made to conform with the RIS assumptions and provides a numerical tool to generate the statistical distributions of interest.

When interdependent potential functions are used, it is not possible to completely separate the conformational partition function and perform straight generalized convolutions. Instead, one can write the conformational energy function (9) as

$$E(\bar{\phi}) = E_1(\phi_1, \ldots, \phi_i) + E_{i+1}(\phi_i, \phi_{i+1}) + E_2(\phi_{i+1}, \ldots, \phi_{n-1}).$$

Then frame densities for the lower and upper segments are generated as before. However, unlike the previously discussed cases, these frame densities are not only functions on G, but also depend on the bond angles contributing to the energy of interaction between the two chain segments. Hence we define f_1 and f_2 by the equalities

$$f_1(g', \phi_i)\mathrm{Vol}(\Delta(g')) = \int_{\{T^{i-1} | g(\phi_1, \ldots, \phi_i) \in \mathrm{Im}(\Delta(g'))\}} \mathrm{e}^{-E_1(\phi_1, \ldots, \phi_i)/k_\mathrm{B}T}\, \mathrm{d}\phi_1 \ldots \mathrm{d}\phi_{i-1}$$

and

$$f_2(g', \phi_{i+1})\mathrm{Vol}(\Delta(g')) = \int_{\{T^{n-i-2} | g(\phi_{i+1}, \ldots, \phi_{n-1}) \in \mathrm{Im}(\Delta(g'))\}} \mathrm{e}^{-E_2(\phi_{i+1}, \ldots, \phi_{n-1})/k_\mathrm{B}T} \times \mathrm{d}\phi_{i+2} \ldots \mathrm{d}\phi_{n-1}.$$

The functions f_1 and f_2 are written more cleanly in the limit of very small voxels using the Dirac delta as

$$f_1(g', \phi_i) = \int_{T^{i-1}} \mathrm{e}^{-E_1(\phi_1, \ldots, \phi_i)/k_\mathrm{B}T} \delta(g^{-1}(\phi_1, \ldots, \phi_i) \circ g') \times \mathrm{d}\phi_1 \ldots \mathrm{d}\phi_{i-1}$$

and

$$f_2(g', \phi_{i+1}) = \int_{T^{n-i-2}} \mathrm{e}^{-E_2(\phi_{i+2}, \ldots, \phi_{n-1})/k_\mathrm{B}T} \times \delta(g^{-1}(\phi_{i+1}, \ldots, \phi_{n-1}) \circ g')\, \mathrm{d}\phi_{i+2} \ldots \mathrm{d}\phi_{n-1}$$

where

$$\delta(g) = \begin{cases} 0 & g \notin \Delta(e) \\ 1/\mathrm{Vol}(\Delta(e)) & g \in \Delta(e) \end{cases}$$

where e is the identity of G (i.e. corresponding to no motion).

These two contributions add to give the composite frame density function using a combination of convolution and weighted integration over the last torsion angle of the first chain segment and first torsion angle of the second chain segment:

$$f(g) = \int_{T^2} \int_G f_1(h, \phi_i) f_2(h^{-1} \circ g, \phi_{i+1}) \times \mathrm{e}^{-E_{i+1}(\phi_i, \phi_{i+1})/k_\mathrm{B}T}\, \mathrm{d}h\, \mathrm{d}\phi_i\, \mathrm{d}\phi_{i+1}.$$

This follows from the evaluation of (7) with (9). From the forms given above for f_1 and f_2, and the properties of the Dirac delta (see Appendix A) it is clear that the normalization

$$\int_G f(g)\, \mathrm{d}g = \int_{T^{n-1}} \mathrm{e}^{-E(\bar{\phi})/k_\mathrm{B}T}\, \mathrm{d}\bar{\phi}$$

holds.

The assumption of discrete sampling of most probable values of torsion angles, as in the RIS model, reduces

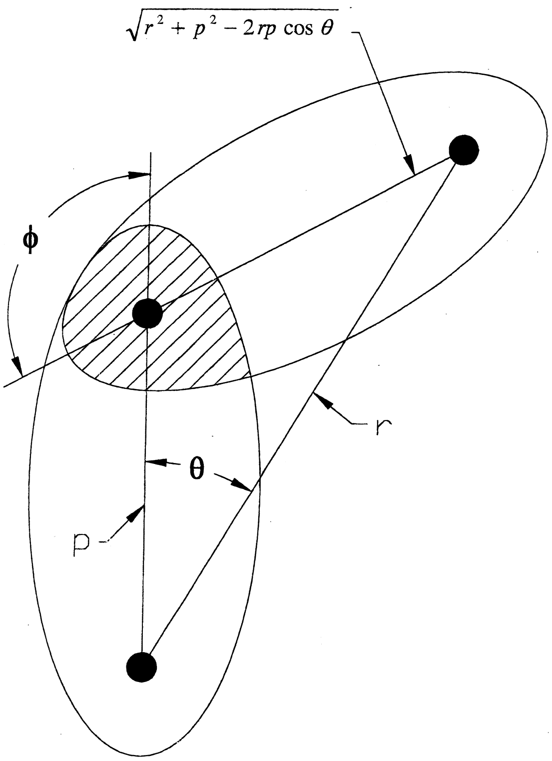

Fig. 2. Approximating excluded volume interactions using $\mu(g,\bar{x})$.

the integrations over T^2 to summations. Hence, from a computational perspective, interdependent energy functions require K^2 convolutions instead of one, where K is the number of sample points for each torsion angle. Typically for an organic chain molecule $K = 3$.

This procedure is iterated in analogy with the algorithm developed in the previous section for the purely kinematic model. Namely, instead of breaking a chain molecule into two imaginary pieces, it can be broken into an arbitrary number of statistically significant segments, and an energy-weighted convolution of each pair of adjacent segments can be performed.

Hence convolution-like integrals of the form

$$f_{i,l}(g,\phi_i,\phi_l) = \int_{T^2}\int_G f_1(h,\phi_i,\phi_k) f_{k+1,l}(h^{-1}\circ g,\phi_{k+1},\phi_l) \times \mathrm{e}^{-E_{i+1}(\phi_k,\phi_{k+1})/k_B T}\,\mathrm{d}h\,\mathrm{d}\phi_k\,\mathrm{d}\phi_{k+1} \qquad (10)$$

are performed, where now both the values of base and distal torsion angles must be recorded so that the process can be iterated. The drawback of this is that if the T^2 integral is approximated as a sum over K^2 values, then the K^2 convolutions on G must be performed for each of the K^2 values of the pairs (ϕ_i,ϕ_l). Hence, K^4 convolutions on G are required at each step. While this is troublesome, we note that these calculations are independent and hence can be distributed over a parallel computer.

4.3. Ensemble properties including long-range conformational energy

In this section we model the long-range interactions in a macroscopically serial chain using an averaging approach which builds on the formulation of the previous sections. The model uses the functions $\mu(g,\bar{x})$ which were not explicitly needed in the purely kinematical model.

Accurately modeling interactions of atoms which are distal in the chain but proximal in space (due to bending of the chain) is one of the most difficult problems in the study of macromolecules, independent of whether they are man-made polymers, proteins, or DNA. Explicitly accounting for all such interactions for all possible configurations by brute force enumeration requires a mind-boggling amount of computational time. At the other extreme, the simplified closed-form analytical models such as the Gaussian random walk do not explicitly account for these interactions. Furthermore, models based on self-avoiding walks and renormalization group methods are respectively limited to rather short chains ($N \ll 100$) or very long ones ($N \longrightarrow \infty$).

A number of different approaches are considered here to incorporate the effects of energy. Perhaps the most straightforward is to penalize contributions in Eq. (6) so that when the support of appropriately shifted functions μ_i and μ_{i+1} intersect, these functions would be disallowed from contributing to the computation of $\mu_{i,i+1}$. Clearly this would generate ensemble statistical distributions which would be lower estimates of those generated from the self-avoiding walk model. A slightly more sophisticated model would calculate the energy of interaction of the distributions μ_i and μ_{i+1} and use this information in an appropriate conformational partition function. We now quantify this discussion. The interaction of segment i and $i+1$ is approximated by considering the interaction of the corresponding functions μ_i and μ_{i+1} as:

$$E_{i,i+1}(h,g) = \int_{\mathbb{R}^3}\int_{\mathbb{R}^3} \mu_i(h,\bar{x})\mu_{i+1}(h^{-1}\circ g, h^{-1}\circ\bar{y}) \times V(\bar{x}-\bar{y})\,\mathrm{d}\bar{x}\,\mathrm{d}\bar{y}. \qquad (11)$$

We assume here that the potential between any two atoms located at positions $\bar{x}$ and $\bar{y}$ is $V(\bar{x}-\bar{y})$. Then Eq. (11) is an approximation of the interaction of all configurations of segment i which terminate at frame h and all configurations of segment $i+1$ with distal end at g and proximal end at h (hence the relative displacement $h^{-1}\circ g$). In the case when the potential function is used to represent pure hard-sphere repulsion and no attraction, then a reasonable model for $V(\cdot)$ is

$$V(\bar{x}) = E_0\delta(\bar{x}).$$

In this case one writes

$$E_{i,i+1}(h,g) = E_0\int_{\mathbb{R}^3} \mu_i(h,\bar{x})\mu_{i+1}(h^{-1}\circ g, h^{-1}\circ\bar{x})\,\mathrm{d}\bar{x}. \qquad (12)$$

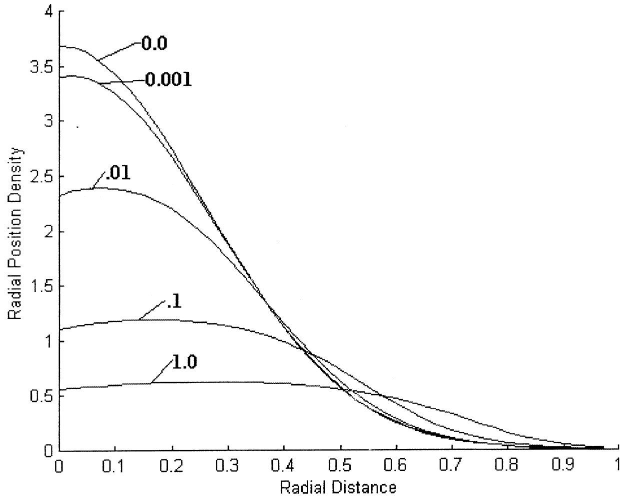

Fig. 3. Energy-weighted PDF for end position.

This may then be used to approximate $\mu_{i,i+1}$ as

$$\mu_{i,i+1}(g,\bar{x}) = \int_G (\mu_i(h,\bar{x}) f_{i+1}(h^{-1}\circ g) + f_i(h)\mu_{i+1}(h^{-1}\circ g, h^{-1}\circ\bar{x}))\, \mathrm{e}^{-E_{i,i+1}(h,g)/\mathrm{k_B}T}\, \mathrm{d}h. \tag{13}$$

By definition, it follows that

$$f_{i,i+1}(g) = \frac{1}{M_i + M_{i+1}} \int_{\mathbb{R}^3} \mu_{i,i+1}(g,\bar{x})\, \mathrm{d}\bar{x} = \int_G f_i(h) f_{i+1}(h^{-1}\circ g)\, \mathrm{e}^{-E_{i,i+1}(h,g)/\mathrm{k_B}T}\, \mathrm{d}h. \tag{14}$$

In this way, the purely kinematical model is modified so as to take into account the energy of interaction of two adjacent segments. As $E_0/\mathrm{k_B}T$ becomes large, the only contributions to $\mu_{i,i+1}$ are from the shifted versions of the functions μ_i and μ_{i+1} which do not overlap at all. This extreme case is a lower bound on the $\mu_{i,i+1}$ that the self avoiding walk would generate. This follows for large $E_0/\mathrm{k_B}T$ because the model in (13) would not only disallow the intersection of two adjacent segments, but also would disallow any contribution if (12) is nonzero, i.e. if μ_i and μ_{i+1} overlap. On the other hand, for smaller values of $E_0/\mathrm{k_B}T$ the present model may provide more realistic statistics than lattice self-avoiding walks since there is no artificial restriction on bond and torsion angles that restrict conformations to a lattice in the present model. In the next section the present formulation is compared with a number of traditional models.

5. Analytical examples

The convolution integral of frame densities for two adjacent segments may be written more explicitly as

$$(f_1 * f_2)(R,\bar{x}) = \int_{\mathrm{SO}(3)} \int_{\mathbb{R}^3} f_1(\mathscr{R},\bar{\xi}) f_2(\mathscr{R}^{\mathrm{T}}R, \mathscr{R}^{\mathrm{T}}(\bar{x}-\bar{\xi})) \times \mathrm{d}\bar{\xi}\, \mathrm{d}\mathscr{R}. \tag{15}$$

In the above equation the integration over G has been rewritten as integration over position and orientation separately, and the volume element is rewritten by observing that it is the product of the volume element for SO(3) and $\mathbb{R}^3$. That is, $\mathrm{d}h = \mathrm{d}\bar{\xi}\mathrm{d}\mathscr{R}$ where $h = (\mathscr{R},\bar{\xi})$, $\mathrm{d}\bar{\xi} = \mathrm{d}\xi_1\, \mathrm{d}\xi_2\, \mathrm{d}\xi_3$, and $\mathrm{d}\mathscr{R}$ is given in Appendix A.

It has been reasoned [12] that for chains with more than a few links, the orientation dependence of the chain PDFs vanishes. That is, to good approximation

$$(f_1 * f_2)(R,\bar{x}) = (f_1 * f_2)(\bar{x}) = \int_{\mathrm{SO}(3)} \int_{\mathbb{R}^3} f_1(\bar{\xi}) \times f_2(\mathscr{R}^{\mathrm{T}}(\bar{x}-\bar{\xi}))\, \mathrm{d}\bar{\xi}\, \mathrm{d}\mathscr{R}.$$

If, in addition, the PDFs f_1 and f_2 are spherically symmetric, so that $f_i(\bar{r}) = f_i(r)$ where $r = |\bar{r}|$, then $f_2(\mathscr{R}^{\mathrm{T}}(\bar{x}-\bar{\xi})) = f_2(\bar{x}-\bar{\xi})$, and so we can calculate the frame density of two adjacent segments using the usual convolution:

$$(f_1 * f_2)(\bar{x}) = \int_{\mathbb{R}^3} f_1(\bar{\xi}) f_2(\bar{x}-\bar{\xi})\, \mathrm{d}\bar{\xi}.$$

5.1. The Gaussian chain

Perhaps the most common model for the distribution of end-vector distribution is the Gaussian distribution:

$$f_{\mathrm{G}}(R,\bar{r})/K^n = \tilde{\rho}'_{\mathrm{G}}(\bar{r}) = \left(\frac{3}{2\pi\langle r^2\rangle}\right)^{3/2} \exp\left[-\frac{3r^2}{2\langle r^2\rangle}\right]. \tag{16}$$

This distribution is spherically symmetric (and hence depends only on $r = |\bar{r}|$). It is normalized so that it is a probability density function,

$$\int_{\mathbb{R}^3} \tilde{\rho}'_{\mathrm{G}}(\bar{r})\, \mathrm{d}r_1\, \mathrm{d}r_2\, \mathrm{d}r_3 = 4\pi \int_0^\infty \tilde{\rho}'_{\mathrm{G}}(r) r^2\, \mathrm{d}r = 1.$$

One observes the equality $\langle r^2\rangle = nl^2$ for a chain with n links each with length l.

It is easy to verify that the PDF for the concatenation of two Gaussian chains is

$$(\tilde{\rho}'_{\mathrm{G}_1} * \tilde{\rho}'_{\mathrm{G}_2})(r) = \left(\frac{3}{2\pi(n_1+n_2)l^2}\right)^{3/2} \exp\left[-\frac{3r^2}{2(n_1+n_2)l^2}\right].$$

That is, the convolution of Gaussians corresponding to segments of length $n_1 l$ and $n_2 l$ results in a Gaussian corresponding to a segment of length $(n_1 + n_2)l$.

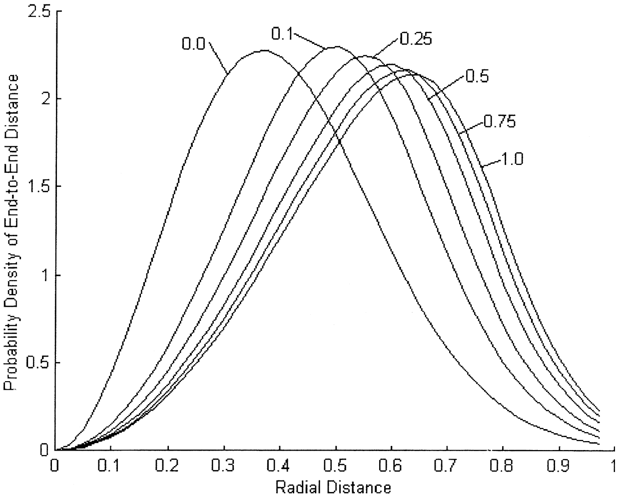

Fig. 4. Energy-weighted end-to-end distance PDF.

5.2. The freely jointed chain

The freely jointed chain model assumes that each link is free to move relative to the others with no constraint on the motion and no correlation between the motion of adjacent links. A derivation of the statistical distribution of end positions of a freely jointed chain can be found in Ref. [6], and is given as

$$\rho_{\mathrm{F}}^{(n)}(\bar{r}) = \frac{1}{2\pi^2 r}\int_0^\infty \sin(qr)[\sin(ql)/ql]^n q\,\mathrm{d}q \qquad (17)$$

where again $r = |\bar{r}|$ and n is the number of links.

Application of the usual Abelian Fourier transform yields

$$\mathscr{F}(\rho_{\mathrm{F}}^{(n)}(\bar{r})) = [\sin(ql)/ql]^n$$

where $q = |\bar{q}|$ and $\bar{q}$ is the vector of Fourier parameters. From this fact it is clear by the classical (Abelian) convolution theorem that

$$\rho_{\mathrm{F}}^{(n_1)} * \rho_{\mathrm{F}}^{(n_2)} = \rho_{\mathrm{F}}^{(n_1+n_2)}.$$

5.3. Self-interacting chains

The previous two examples illustrate that the generalized convolution model agrees with classical formulations in the degenerate case of a spherically symmetric density function without dependence on orientation when the effects of excluded volume are not included. We now show how this formulation can be used to approximate the effects of excluded volume using an analytically defined function μ.

To begin, let us assume that segment i is long enough to be considered completely flexible (though inextensible). Then the function $\mu_i(g, \bar{x})$ will only depend on the magnitude of the translation $r = |\bar{r}|$ where $g = (R, \bar{r})$. If the length of this segment is L_i, then for given r, all configurations of the segment will be contained within an axially symmetric ellipsoid with axis in the direction of $\bar{r}$, and foci at the proximal and distal ends of the segment. For the moment let us assume that $\bar{r} = [r, 0, 0]^{\mathrm{T}}$. Then this ellipsoid will be defined by the equation

$$D_i(r, \bar{x}) = \left(\frac{x - r/2}{a_i}\right)^2 + \frac{y^2 + z^2}{b_i^2(r)} - 1 = 0.$$

The parameters a_i and $b_i(r)$ are derived from r and L_i as

$$b_i(r) = \frac{1}{2}\sqrt{L_i^2 - r^2}, \qquad a_i = L_i/2.$$

A corresponding model for the function $\mu_i(g, \bar{x})$ is

$$\mu_i(r, \bar{x}) = \frac{M_i f_i(r)}{\frac{4}{3}\pi a_i b_i^2(r)}\mathrm{step}[-D_i(r, \bar{x})].$$

Here step[x] is the unit Heaviside step function which takes the value 1 for $x \geq 0$ and zero otherwise. The denominator in the fraction above is the volume of an ellipsoid of revolution defined by major axis with length a_i and minor axes with length b_i. Hence this μ_i is constant over the interior of the ellipsoid, zero outside of it, and normalized so that $\int_{\mathbb{R}^3} \mu_i \,\mathrm{d}\bar{x} = M_i f_i$.

The energy of interaction between the ensembles for two adjacent segments is then of the form

$$\begin{aligned} &E_{i,i+1}(\phi, r_1, r_2) \\ &\quad = \frac{E_0 M_i M_{i+1} f_i(r_1) f_{i+1}(r_2)}{\frac{4}{3}\pi a_i b_i^2(r_1)\cdot \frac{4}{3}\pi a_{i+1} b_{i+1}^2(r_2)}\int_{\mathbb{R}^3} \mathrm{step}[-D_i(r_1, \bar{x})] \\ &\qquad \times \mathrm{step}[-D_{i+1}(r_2, R_z(-\phi)\bar{x})]\,\mathrm{d}\bar{x}. \end{aligned}$$

The integral is simply the volume of intersection of two ellipsoids which share one of their foci and are rotated relative to each other by angle ϕ. Clearly this energy function is 2π-periodic in ϕ and has the symmetries

$$E_{i,i+1}(\phi, r_1, r_2) = E_{i,i+1}(-\phi, r_1, r_2) = E_{i,i+1}(\phi, r_2, r_1).$$

Since we have assumed that $f_i(g) = f_i(r)$, Eq. (14) reduces to

$$\begin{aligned} f_{i,i-1}(r) = 2\pi &\int_0^\infty f_i(p)\int_0^\pi f_{i+1}\left(\sqrt{r^2 + p^2 - 2rp\cos\theta}\right) \\ &\times \exp\left(\frac{-1}{k_B T}E_{i,i+1}\left(\phi(r, p, \theta), p, \sqrt{r^2 + p^2 - 2rp\cos\theta}\right)\right) \\ &\times p^2\,\mathrm{d}p\,\mathrm{d}\theta. \end{aligned} \qquad (18)$$

The functional relationship between ϕ and the variables r, p, and θ is easily derived from Fig. 2 using trigonometry as

$$\sin\phi = \frac{r\sin\theta}{\sqrt{r^2 + p^2 - 2rp\cos\theta}},$$

$$\cos\phi = \frac{p - r\cos\theta}{\sqrt{r^2 + p^2 - 2rp\cos\theta}}.$$

These are inverted using the Atan2( ) function to yield $\phi(r,p,\theta)$.

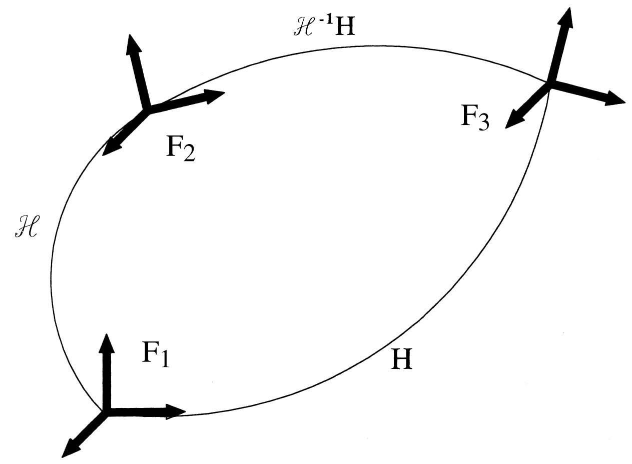

Fig. 5. Concatenation of homogeneous transformations.

As a numerical example, we take f_i and f_{i+1} to be Gaussian with $n_i = n_{i+1} = 10$ and $l = 0.05$. Substituting into Eq. (18), the resulting $f_{i,i+1}$ (normalized to be a probability density function) is shown in Fig. 3 for the values of $\epsilon = M_i M_{i+1} E_0/\mathrm{k_B}T$ indicated. Fig. 4 plots the end-to-end distance function d(r) defined in Eq. (3) for the values of ϵ shown. These plots agree with the qualitative sense that the mean end-to-end distance should increase with the degree of repulsion.

6. Conclusions

The ensemble properties of chain macromolecules are generated by convolving functions of rigid-body motion. Analytical examples demonstrate the technique for both the case when interactions between distal segments in the chain are modeled and when they are not. A general numerical scheme to implement the analytical formulation is outlined, and the complexity of the approach is analyzed. This computational approach is very general, and can be made to conform to a variety of steric constraints and conformational energy functions. Analytical examples show that the present formulation includes classical models such as the Gaussian and freely jointed chains. The present work also shows how the PDF of end positions and orientations for the RIS model as well as models including the long-range effects of excluded volume can be generated. Finally, it is worth noting that while the formulation presented here is for PDFs defined on continuous domains, the restriction of chains to lattices, and the use of the corresponding discrete crystallographic motions in place of continuous rigid-body motions may lead to even faster numerical implementations.

Acknowledgements

Thanks go to D. Healy, G. Domokos and M. Donohue for each independently suggesting that mathematical techniques originally developed by the author for the field of robotics be applied in the polymer context. The work that gave rise to this was supported under NSF grants NSF-IRI/RHA 97-31720 and a 1994 Presidential Faculty Fellow Award. D. Stein is acknowledged for generating the figures in this paper.

Appendix A. Integration and convolution of motion-dependent functions

An arbitrary rigid-body motion can be viewed as the pair $g = (R, \bar{r})$ where $R \in \mathrm{SO}(3)$ (i.e. R is a 3×3 rotation matrix), and $\bar{r} \in \mathbb{R}^3$ is a translation vector in three-dimensional space. The composition law is $g_1 \circ g_2 = (R_1 R_2, R_1 \bar{r}_2 + \bar{r}_1)$. The action of the motion g on a position vector $\bar{x} \in \mathbb{R}^3$ is $g \circ \bar{x} = R\bar{x} + \bar{r}$.

The collection of all rigid-body motions is denoted in this paper as G. Any $g \in G$ can be faithfully represented with a 4×4 *homogeneous transformation matrix* of the form:

$$H(g) = \begin{pmatrix} R & \bar{r} \\ \bar{0}^{\mathrm{T}} & 1 \end{pmatrix}.$$

Henceforth no distinction is made between G and the set of all 4×4 homogeneous transformation matrices.

A.1. Invariant integration

Orientations (or rotations) in three-dimensional space are parameterized with *ZXZ* Euler angles (Flory calls these the “Eulerian Angles” [6, p. 404])

$$R(\phi, \theta, \psi) = R_z(\phi) R_x(\theta) R_z(\psi)$$

$$= \begin{pmatrix} \mathrm{c}\psi\mathrm{c}\phi - \mathrm{c}\theta\mathrm{s}\phi\mathrm{s}\psi & -\mathrm{s}\psi\mathrm{c}\phi - \mathrm{c}\theta\mathrm{s}\phi\mathrm{c}\psi & \mathrm{s}\theta\mathrm{s}\phi \\ \mathrm{c}\psi\mathrm{s}\phi + \mathrm{c}\theta\mathrm{c}\phi\mathrm{s}\psi & -\mathrm{s}\psi\mathrm{s}\phi + \mathrm{c}\theta\mathrm{c}\phi\mathrm{c}\psi & -\mathrm{s}\theta\mathrm{c}\phi \\ \mathrm{s}\theta\mathrm{s}\psi & \mathrm{s}\theta\mathrm{c}\phi & \mathrm{c}\theta \end{pmatrix}$$

where ‘s’ is shorthand for ‘sin’ and ‘c’ is shorthand for ‘cos’. The range of these angles is $0 \le \phi, \psi \le 2\pi$ and $0 \le \theta \le \pi$. The volume element for G is given by

$$\mathrm{d}g = \frac{1}{8\pi^2} \sin\theta \,\mathrm{d}\phi\,\mathrm{d}\theta\,\mathrm{d}\psi\,\mathrm{d}r_1\,\mathrm{d}r_2\,\mathrm{d}r_3$$

which is the product of the volume elements for $\mathbb{R}^3$ ($\mathrm{d}\bar{r} = \mathrm{d}r_1 \mathrm{d}r_2 \mathrm{d}r_3$), and for SO(3) ($\mathrm{d}R = (1/8\pi^2) \sin\theta \,\mathrm{d}\phi\,\mathrm{d}\theta\,\mathrm{d}\psi$). The normalization factor in the definition of dR is so that $\int_{\mathrm{SO}(3)} \mathrm{d}R = 1$.

The fact that this volume element is invariant to right and left translations, i.e.

$$\mathrm{d}g = \mathrm{d}(h \circ g) = \mathrm{d}(g \circ h)$$

is well known in certain communities (see, e.g. Refs. [32,33]).

We note the following shorthands used throughout the paper:

$$\int_{SO(3)} = \int_{\phi=0}^{2\pi} \int_{\theta=0}^{\pi} \int_{\psi=0}^{2\pi};$$

$$\int_{\mathbb{R}^3} = \int_{r_1=-\infty}^{\infty} \int_{r_2=-\infty}^{\infty} \int_{r_3=-\infty}^{\infty};$$

and

$$\int_G = \int_{\mathbb{R}^3} \int_{SO(3)}.$$

A.2. Geometrical interpretation of convolution on G

Suppose there are three frames in space, F_1, F_2 and F_3, as shown in Fig. 5. The first frame can be viewed as fixed, the second frame as moving with respect to the first, and the third frame as moving with respect to the second. Let the homogeneous transform $\mathscr{H}$ describe the position and orientation of F_2 w.r.t. F_1, and H' describe the position and orientation of F_3 w.r.t. F_2. Then the position and orientation of F_3 w.r.t. F_1 is $H = \mathscr{H}H'$. The position and orientation of F_3 w.r.t. F_2 can then be written as

$$H' = \mathscr{H}^{-1}H.$$

We may divide up G into volume elements, or “voxels,” of finite but small size. The volume of the voxel centered at $H \in G$ is denoted ΔH, and as the element size is chosen smaller and smaller it becomes closer to the differential volume element dH.

The motion of F_2 relative to F_1 and the motion of F_3 relative to F_2 can both be considered elements of G, and no distinction is made between these motions and the transformation matrices $\mathscr{H}$ and H' which represent these motions.

Assuming we move $\mathscr{H}$ and H' through a finite number of different positions and orientations, let f_1 be a function that records how often the $\mathscr{H}$ frames appear in each voxel, divided by the voxel volume $\Delta\mathscr{H}$. Likewise, let f_2 be the function describing how often the H' frames appear in each voxel normalized by voxel volume.

To calculate how often the H frames appear in each voxel in G for all possible values of $\mathscr{H}$ and H', we may perform the following steps:

- Evaluate $f_1 = f_1(\mathscr{H})$ (frequency of occurrence of $\mathscr{H}$).
- Evaluate $f_2 = f_2(H') = f_2(\mathscr{H}^{-1}H)$ frequency of occurrences of $H' = \mathscr{H}^{-1}H$).
- Weight (multiply) the left-shifted histogram, $f_2(\mathscr{H}^{-1}H)$, by the number of frames which are doing the shifting. This number is $f_1(\mathscr{H})\Delta\mathscr{H}$ for each $\mathscr{H}$.[4]
- Sum (integrate) over all these contributions:

$$(f_1 * f_2)(H) = \int_G f_1(\mathscr{H}) f_2(\mathscr{H}^{-1}H)\, d\mathscr{H}.$$

Using this definition of convolution, the Dirac δ-function defined in the text has the property

$$\delta * f = f * \delta = f$$

for any well-behaved $f(g)$.

References

[1] Kuhn W. Kolloid-Z 1934;68:2.
[2] Flory PJ. J Chem Phys 1949;17(3):303–10.
[3] Volkenstein MV. Conformational statistics of polymeric chains, New York: Interscience, 1963 originally published in Russian by: IzdR-tel'stvo Akad. Nauk SSSR, Moscow, 1959.
[4] Volkenstein MV. Dokl Akad Nauk SSSR 1951;78:879.
[5] Volkenstein MV, Ptitsyn OB. Dokl Akad Nauk SSSR 1953;91:1313.
[6] Flory PJ. Statistical mechanics of chain molecules, New York: Wiley-Interscience, 1969.
[7] Flory PJ. Macromolecules 1974;7(3):381.
[8] Birshtein TM, Ptitsyn OB. Conformations of macromolecules, New York: Interscience, 1966.
[9] Mattice WL, Suter UW. Conformational theory of large molecules, the rotational isomeric state model in macromolecular systems, New York: Wiley, 1994.
[10] Freed KF. Renormalization group theory of macromolecules, New York: Wiley, 1987.
[11] des Cloizeaux J, Jannink G. Polymers in solution: their modeling and structure, Oxford: Clarendon Press, 1990.
[12] de Gennes PG. Introduction to polymer dynamics, Cambridge: Cambridge University Press, 1990.
[13] de Gennes PG. Scaling concepts in polymer physics, Ithaca, NY: Cornell University Press, 1979.
[14] Doi M, Edwards SF. The theory of polymer dynamics, Oxford: Clarendon Press, 1986.
[15] Yamakawa H. Helical wormlike chains in polymer solutions, Berlin: Springer, 1997.
[16] Lennard-Jones JE. Physica 1937;4:941.
[17] Rosenbluth MN, Rosenbluth AW. J Chem Phys 1955;23(2):356–9.
[18] Alexandrowicz Z. J Chem Phys 1969;51(2):561–5.
[19] Collet O, Premilat S. Macromolecules 1993;26:6076–80.
[20] Madras N, Slade G. The self-avoiding walk, Boston: Birkhauser, 1996.
[21] Norisuye T, Tsuboi A, Teramoto A. Polym J 1996;28(4):357–61.
[22] Mondescu RP, Muthukumar M. Phys Rev E 1998;57(4):4411–9.
[23] He S, Scheraga HA. J Chem Phys 1998;108(1):271–86.
[24] Plimpton S, Hendrickson B. J Comput Chem 1996;17(3):326–37.
[25] Grosberg AYu, Khokhlov AR. Statistical physics of macromolecules, New York: American Institute of Physics, 1994.
[26] Nechaev S, Desbois J. J Stat Phys 1997;88(1/2):201–29.
[27] Kloczkowski A, Jernigan RL. Comput Theoret Polym Sci 1997;7(3/4):163–73.
[28] Honeycutt JD. Comput Theoret Polym Sci 1998;8(1/2):1–8.
[29] Thirumalai D, Ha B-Y. In: Grosberg A, editor. Theoretical and mathematical models in polymer research, New York: Academic Press, 199. pp. 1–35.
[30] Chirikjian GS. Inverse Problems 1996;12:579–99.
[31] Chirikjian GS, Ebert-Uphoff I. IEEE Trans Robotics Autom 1998;14(1):123–36.
[32] Sugiura M. Unitary representations and harmonic analysis, 2. Amsterdam: North Holland, 1990.
[33] Vilenkin NJ, Klimyk AU. Representation of lie groups and special functions, 1–3. Dordrecht: Kluwer Academic, 1991.