Random walks with non-Gaussian step-size distributions and the folding of random polymer chains

R. H. A. David Shaw* and J. A. Tuszyński†
Department of Physics, University of Alberta, Edmonton, Alberta, Canada T6G 2J1
(Received 29 August 2002; published 13 March 2003)

In this paper, we study a random walker whose step-size distribution is of non-Gaussian bimodal form due to the addition of a quartic term in the exponential. By the central limit theorem, we know that in the limit of a large number of steps, the probability distribution representing the distance the walker has traveled becomes Gaussian. We investigate the nature of this convergence both numerically and analytically. We obtain a scaling relation describing the number of steps required for convergence in terms of the width and separation of the peaks of the step-size distribution. We assume in the concluding section that our model is well suited for the application of the folding of a random polymer chain.

DOI: 10.1103/PhysRevE.67.031102 PACS number(s): 05.40.−a, 05.20.−y, 02.50.−r

I. INTRODUCTION

The random walk problem has a long and illustrious history in statistical physics [1]. Applications of the random walk concept include not only the diffusion of molecules in a gas [2] and colloidal suspensions [1] but also spin-$\frac{1}{2}$ paramagnetism, light intensity due to N incoherent light sources [2], and even spatial distribution of stars [1], to name but a few. Recently, a resurgence of interest in random walks has been experienced as a result of applications to fractal media and percolation phenomena [3].

One remarkable feature of a large subset of random walk processes, predicted by the central limit theorem, is that given a sufficiently large number of steps, the probability distribution representing the location of a random walker will become arbitrarily close to a Gaussian form [4]. It has remained unclear, however, exactly how this convergence will occur over time for a general step-size distribution. Indeed, for step-size distributions with strongly bimodal characteristics, how such convergence will occur is an intriguing and nontrivial question.

In this paper, we intend to determine the nature of the convergence for a particular choice of step-size distribution—a Gaussian with an anharmonic quartic term in the exponential. Both analytic and numerical methods of solution will be employed; we will also introduce an approximation which is valid over a large range of parameter space. Additionally, the limit as one approaches a binomial distribution will be examined.

The motivation to investigate the issue of a bimodal (but not necessarily deltoidal) step-size distribution comes from several areas. First, the probability distribution that we study here appears in the so-called s^4, one-dimensional, continuous Ising model [5]. Bimodal probability distributions have also been discussed in connection with the Gaussian ensemble as an interpolating ensemble in finite-size systems exhibiting analogs of critical phenomena [6]. Indeed, in the context of criticality, the step-size distribution $w(s)$ discussed later in the paper acquires a different meaning than that commonly associated with random walks. The stochastic variable s may be interpreted as the size of a given domain of a thermodynamic phase in a multistable system. The issue one would like to address is the transition from a local double-peaked probability distribution to a global single-peaked distribution as the number of steps (size of the system) is increased to infinity.

It is interesting to note that there is a connection between the random walk problem and the Fokker-Planck equation for the probability distribution [1]. When the random walk process takes place in the presence of a bistable quartic potential of the type $V(x) = -x^2/2 + x^4/4$, then the Fokker-Planck equation for the probability distribution $P(x,t)$ takes the form

$$\frac{\partial P}{\partial t} = -\frac{\partial}{\partial x}[(x-x^3)P] + D\frac{\partial^2 P}{\partial x^2}, \tag{1.1}$$

and its stationary solution is a double-peaked quartic non-Gaussian function [7]

$$P_{st} = C \exp\left[-\frac{(x^2-1)^2}{4D}\right], \tag{1.2}$$

which will play a central role in this paper. A similar problem arose in connection with the threshold for laser action [8].

Finally, it is worth mentioning that a practical reason for studying a random walker with a "diffuse" step-size distribution may relate to the behavior of motor proteins [9,10] such as myosin, dynein, and kinesin, which bind to and unbind from biopolymer filaments such as actin and microtubules. The underlying periodicity of the substrate potential for the motion of these Brownian motors may be associated with a step size, which is not sharply defined.

II. RANDOM WALKS WITH A STEP-SIZE DISTRIBUTION

The probability $P(x)$ that after N steps a random walker has reached the location x given a step-size distribution $w(s)$, which remains the same for each step, is expressed in

*Electronic address: dshaw@phys.ualberta.ca
†Electronic address: jtus@phys.ualberta.ca

terms of an N-dimensional integral [2]:

$$P(x)=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}w(s_1)w(s_2)\cdots w(s_N)\delta\times\left(\sum_{i=1}^{N}s_i-x\right)ds_1ds_2\cdots ds_N. \tag{2.1}$$

The problem can be simplified somewhat by Fourier transforming to "momentum" space, allowing us to represent the probability distribution as follows:

$$P(x)=\frac{1}{2\pi}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}\exp\left[ik\left(\sum_{j=1}^{N}s_j-x\right)\right]\times w(s_1)w(s_2)\cdots w(s_N)dkds_1ds_2\cdots ds_N. \tag{2.2}$$

We can now write the bulk of the integral as a product,

$$P(x)=\frac{1}{2\pi}\int_{-\infty}^{\infty}dke^{-ikx}\prod_{j=1}^{N}\int_{-\infty}^{\infty}e^{iks_j}w(s_j)ds_j, \tag{2.3}$$

which, with the definition of $Q(k)$ as the Fourier transform of the step-size distribution $w(s)$,

$$Q(k)=\int_{-\infty}^{\infty}e^{iks}w(s)ds, \tag{2.4}$$

permits the further simplification of Eq. (2.3), giving

$$P(x)=\frac{1}{2\pi}\int_{-\infty}^{\infty}e^{-ikx}Q^N(k)dk. \tag{2.5}$$

The expression in Eq. (2.5) is the most general form for $P(x)$. However, as is readily seen from Eq. (2.1), in the special case $N=2$, it is possible to write $P(x)$ as a simple convolution,

$$P(x)=\int_{-\infty}^{\infty}w(x-s)w(s)ds, \tag{2.6}$$

which, as will be seen later, leads to some interesting three-dimensional plots that are useful in illustrating the nature of the problem.

Using the central limit theorem [4], it can be shown that for a wide range of step-size distributions $w(s)$, $P(x)$ approaches a Gaussian form in the limit of large N. A clear illustration of this result can be found in Ref. [2], based on the assumptions that each step is statistically independent and that $|w(s)|\to 0$ sufficiently fast as $|s|\to\infty$. For small values of k, expanding e^{iks} in a Taylor series yields

$$Q(k)=\sum_{n=0}^{\infty}\frac{(ik)^n}{n!}\int_{-\infty}^{\infty}s^nw(s)ds. \tag{2.7}$$

Recalling the definition of the nth moment of s,

$$\overline{s^n}=\int_{-\infty}^{\infty}s^nw(s)ds, \tag{2.8}$$

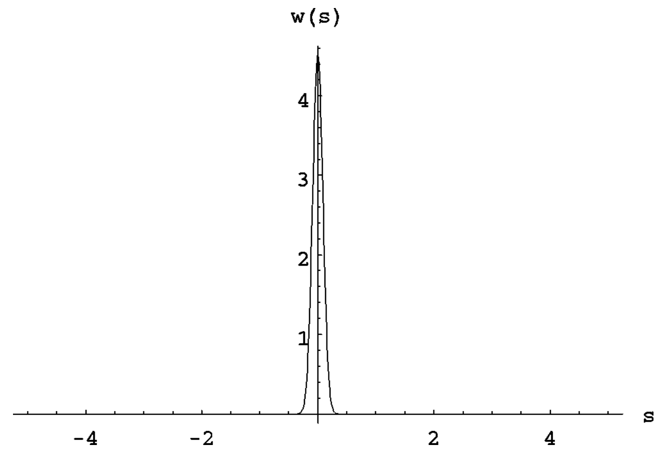

FIG. 1. Step-size distribution $w(s)$ for the case $\beta=-8$ and $\gamma=2$.

we write

$$Q(k)=\sum_{n=0}^{\infty}\frac{(ik)^n}{n!}\overline{s^n}. \tag{2.9}$$

Using the expansion in Eq. (2.9) above, we have

$$\ln Q^N(k)=N\ln\left[\sum_{n=0}^{\infty}\frac{(ik)^n}{n!}\overline{s^n}\right]. \tag{2.10}$$

Now, using the Taylor expansion of the logarithm, we obtain

$$\ln Q^N(k)\approx N\ln\left[i\bar{s}k-\frac{1}{2}\overline{s^2}k^2-\frac{1}{2}(i\bar{s}k)^2\right] \tag{2.11}$$

which, recalling the definition of the variance $\overline{(\Delta s)^2}=\overline{s^2}-\bar{s}^2$ gives us

$$\ln Q^N(k)\approx N\ln\left[i\bar{s}k-\frac{1}{2}\overline{(\Delta s)^2}k^2\right]. \tag{2.12}$$

Therefore, exponentiating the above and inserting into Eq. (2.5) yields

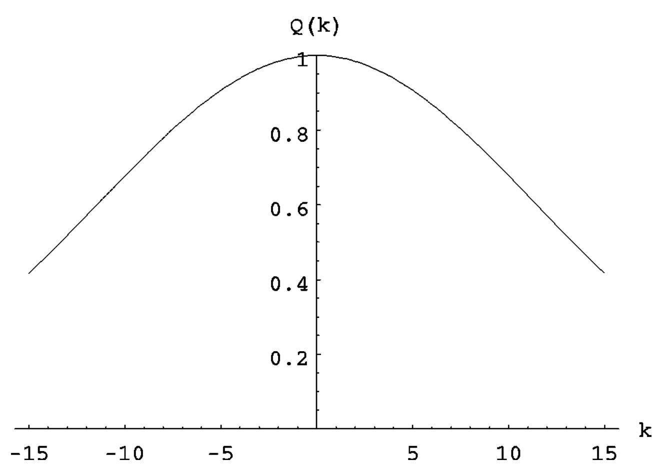

FIG. 2. Fourier transform $Q(k)$ for the case $\beta=-8$ and $\gamma=2$.

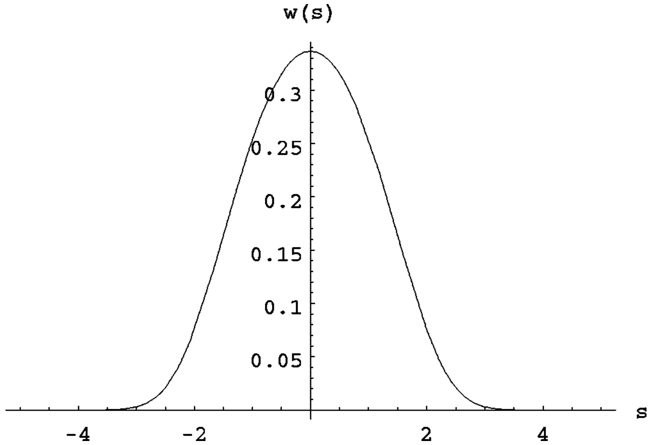

FIG. 3. Step-size distribution $w(s)$ for the case $\beta=-\frac{1}{32}$ and $\gamma=2$.

$$P(x)=\frac{1}{2\pi}\int_{-\infty}^{\infty}e^{i(N\overline{s}-x)k-(1/2)N\overline{(\Delta s)^2}k^2}dk. \qquad (2.13)$$

This integral can be easily performed, giving

$$P(x)=\frac{1}{\sqrt{2\pi N\overline{(\Delta s)^2}}}e^{-(x-N\overline{s})^2/2N\overline{(\Delta s)^2}}, \qquad (2.14)$$

which is exactly the form of a Gaussian distribution, with mean μ, and variance σ^2 given by

$$\mu=N\overline{s}; \quad \sigma^2=N\overline{(\Delta s)^2}. \qquad (2.15)$$

We note, however, that in all of the above, it has not been demonstrated how the convergence takes place. This intermediate regime, between $N=1$ and N large, is our area of interest in this paper.

We wish to consider a Gaussian distribution modified by the addition of an anharmonic quartic term in the exponential. This corresponds to

$$w(s)=\frac{1}{w_0}e^{-\alpha s^2-\beta s^4}. \qquad (2.16)$$

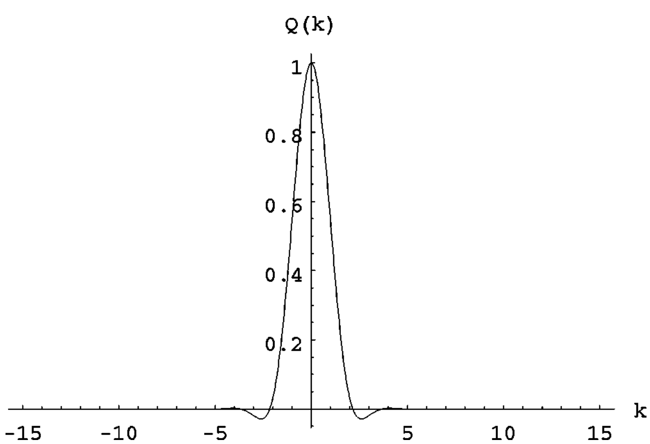

FIG. 4. Fourier transform $Q(k)$ for the case $\beta=-\frac{1}{32}$ and $\gamma=2$.

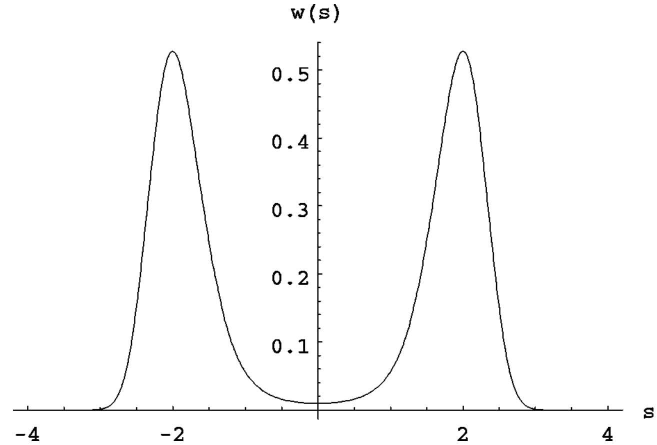

FIG. 5. Step-size distribution $w(s)$ for the case $\beta=\frac{1}{4}$ and $\gamma=2$.

Functions of this form are of interest in a number of distinct areas of study in physics, in which the bimodality of the system is an important characteristic—the Landau-Ginzburg model of phase transitions and the stochastic development of a bifurcation [11] are two examples. It is more useful to represent $w(s)$ as

$$w(s)=\frac{1}{w_0}e^{2\beta\gamma^2 s^2-|\beta|s^4}, \qquad (2.17)$$

where $\beta\in(-\infty,\infty)$ and γ represents the location of the peaks in the case $\beta<0$ (see Figs. 1–8). w_0 is the normalization factor, and is calculated to be

$$w_0=\begin{cases}\dfrac{\pi\gamma}{2}e^{-\beta\gamma^4/2}\left[I_{-1/4}\left(\dfrac{\beta\gamma^4}{2}\right)+I_{1/4}\left(\dfrac{\beta\gamma^4}{2}\right)\right], & \beta>0\\ \dfrac{\gamma}{\sqrt{2}}e^{-|\beta|\gamma^4/2}K_{1/4}(|\beta|\gamma^4), & \beta<0.\end{cases} \qquad (2.18)$$

The above result can be stated more succinctly using parabolic cylinder functions $D_\nu(x)$ of order $-\frac{1}{2}$ (see Appendix A), but we use the more familiar modified Bessel functions for ease of visualization [12].

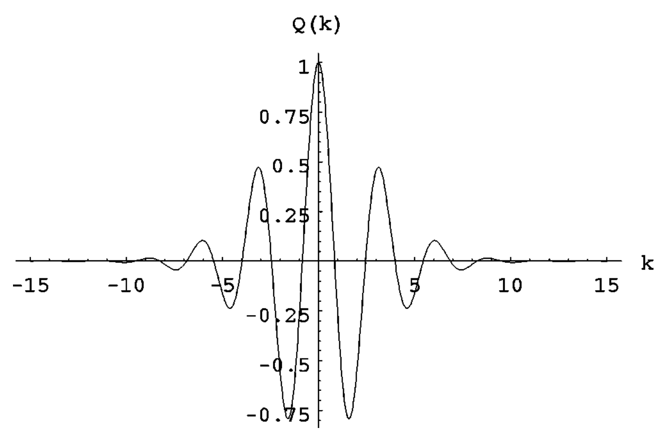

FIG. 6. Fourier transform $Q(k)$ for the case $\beta=\frac{1}{4}$ and $\gamma=2$.

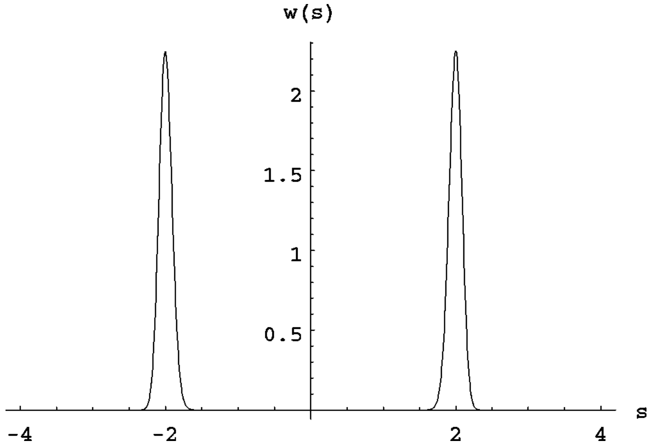

FIG. 7. Step-size distribution $w(s)$ for the case $\beta=4$ and $\gamma=2$.

Finding $Q(k)$ corresponds to computing the Fourier transform of our non-Gaussian step-size distribution $w(s)$,

$$Q(k)=\frac{1}{w_0}\int_{-\infty}^{\infty}e^{iks+2\beta\gamma^2s^2-|\beta|s^4}ds. \tag{2.19}$$

To the best of our knowledge, a closed form analytic solution for this integral does not exist. However, if one expands e^{iks} in a power series, one may express $Q(k)$ as an infinite series of integrals as follows:

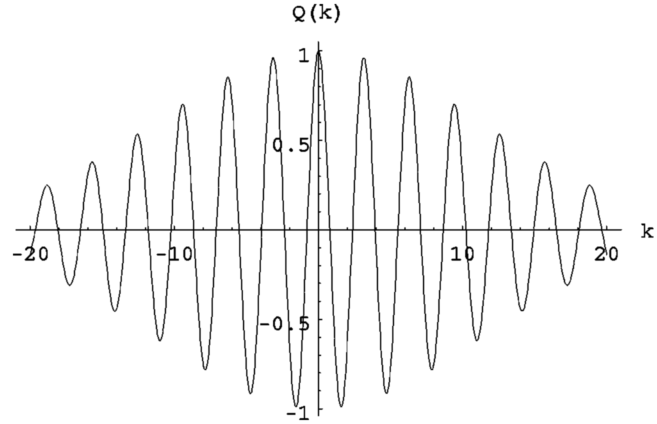

FIG. 8. Fourier transform $Q(k)$ for the case $\beta=4$ and $\gamma=2$.

$$Q(k)=\frac{1}{w_0}\sum_{n=0}^{\infty}\frac{(-1)^nk^{2n}}{(2n)!}\int_{-\infty}^{\infty}s^{2n}e^{2\beta\gamma^2s^2-|\beta|s^4}ds. \tag{2.20}$$

Each individual integral is expressible in closed form, and leads to an infinite sum of hypergeometric functions [13] (see Appendix B) as follows:

$$Q(k)=\sum_{n=0}^{\infty}\frac{(-1)^nk^{2n}}{w_0(2n)!}\times\begin{cases}\frac{\pi}{2}\beta^{-1/4-n/2}\left[\Gamma\left(\frac{1}{4}+\frac{n}{2}\right){}_1F_1\left(\frac{1}{4}+\frac{n}{2},\frac{1}{2},\beta\gamma^4\right)+2\sqrt{\beta}\gamma^2\Gamma\left(\frac{3}{4}+\frac{n}{2}\right){}_1F_1\left(\frac{3}{4}+\frac{n}{2},\frac{1}{2},\beta\gamma^4\right)\right], & \beta>0\\ \frac{1}{2}|\beta|^{-3/4-n/2}\left[\sqrt{|\beta|}\Gamma\left(\frac{1}{4}+\frac{n}{2}\right){}_1F_1\left(\frac{1}{4}+\frac{n}{2},\frac{1}{2},\frac{|\beta|\gamma^4}{2}\right)+2\beta\gamma^2\Gamma\left(\frac{3}{4}+\frac{n}{2}\right){}_1F_1\left(\frac{3}{4}+\frac{n}{2},\frac{1}{2},\frac{|\beta|\gamma^4}{2}\right)\right], & \beta<0.\end{cases} \tag{2.21}$$

Unfortunately, these series are alternating, and, in general, require about 50 to 100 terms to converge. Moreover, the magnitude of the first several terms is rather large (of the order of a trillion or so), which can introduce significant roundoff errors if one attempts to sum the series directly without taking appropriate precautions.

III. NUMERICAL INTEGRATION RESULTS

Ideally, one would like an analytic solution for $P(x)$, to allow one to study in detail the convergence to a Gaussian distribution. This requires, in turn, an analytic solution for $Q(k)$. While our series expansion for $Q(k)$ in Eq. (2.21) is indeed analytic, it does not admit a simple closed form solution for $P(x)$.

Numerical integration of the series representation for $Q(k)$ is possible, but is a rather lengthy and difficult procedure. One must evaluate the sum of the first 100 terms of the series at each point in the numerical integration algorithm, keeping 20 digits precision to avoid the roundoff error noted in the preceding section. Such computations are very time and resource intensive. Other numerical methods allow for faster computation of $Q(k)$, but at the same time, lead to numerical difficulties in determining $P(x)$. Numerical integration of a numerically obtained function is a rather involved process. In particular, for the case where $\beta>0$, $Q(k)$ is a rapidly oscillating function, further increasing the difficulty of computation. In the special case $N=2$, as we have seen in Eq. (2.6), $P(x)$ can be expressed as a simple convolution. Although these integrals have, in general, no analytic solution, it is possible to evaluate them numerically. Unfortunately, attempts to solve cases for larger N by analogous methods rapidly become too difficult to carry out numerically. Studies of behavior for large N are, therefore, ruled out. Almost all methods used to date result, at best, in a numerical solution for $P(x)$. While this allows for visualiza-

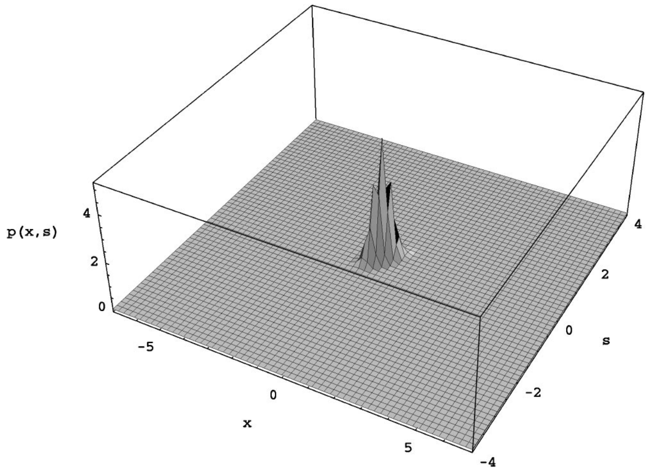

FIG. 9. Structure function $p(x,s)$ for the case $\beta=-2$ and $\gamma=2$.

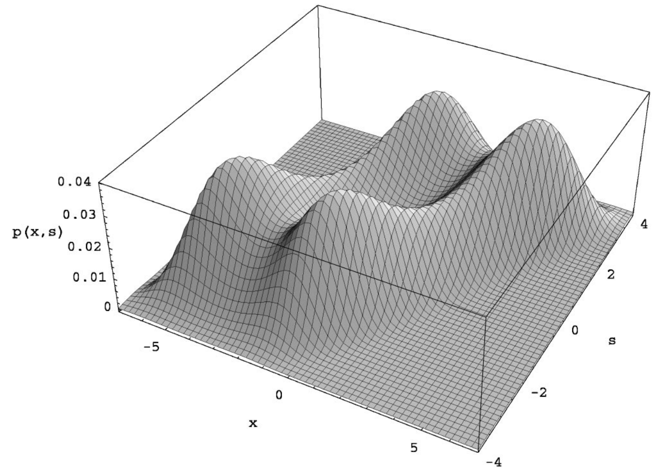

FIG. 11. Structure function $p(x,s)$ for the case $\beta=\frac{1}{32}$ and $\gamma=2$.

tion of the functional form, it unfortunately does not easily admit study of the dependence of the convergence to a Gaussian form on the parameters N, β, or γ. Without a simpler analytic form for $Q(k)$, we have little hope of obtaining an exact analytic form—and the corresponding insight that would bring—for $P(x)$.

For the purpose of illustration, in Figs. 1–8, we plot four cases of the step-size distribution $w(s)$ and their corresponding Fourier transforms $Q(k)$. These cases are typical of the different types of behavior which one may encounter in the course of this problem.

Plotting the integrand of the convolution in Eq. (2.6) is useful in visualizing how various aspects of $P(x)$ arise. Four examples illustrating some characteristic behaviors of this ‘‘structure function,’’ $p(x,s)$, are plotted in Figs. 9–12.

IV. USING A CONVENIENT APPROXIMATION

As it became clear in Sec. III, one must find some approximation for $w(s)$ with a simple $Q(k)$ in order to be able to proceed further with this problem. For the case, where $\beta\gamma^2\geqslant 8$, $w(s)$ is very well approximated by two Gaussians centred at $\pm\gamma$, respectively. The appropriate functional form is

$$\tilde{w}(s)=\sqrt{\frac{\beta\gamma^2}{\pi}}[e^{-4\beta\gamma^2(s-\gamma)^2}+e^{-4\beta\gamma^2(s+\gamma)^2}], \quad (4.1)$$

which has an easily computable Fourier transform, namely,

$$\tilde{Q}(k)=e^{-k^2/16\beta\gamma^2}\cos(k\gamma). \quad (4.2)$$

The exponentiation of $\tilde{Q}(k)$ is straightforward, yielding

$$\tilde{Q}^N(k)=e^{-Nk^2/16\beta\gamma^2}\cos^N(k\gamma). \quad (4.3)$$

Using Gradshteyn and Ryzhik [12], we express powers of cosine as a sum of cosine functions with multiple angle arguments

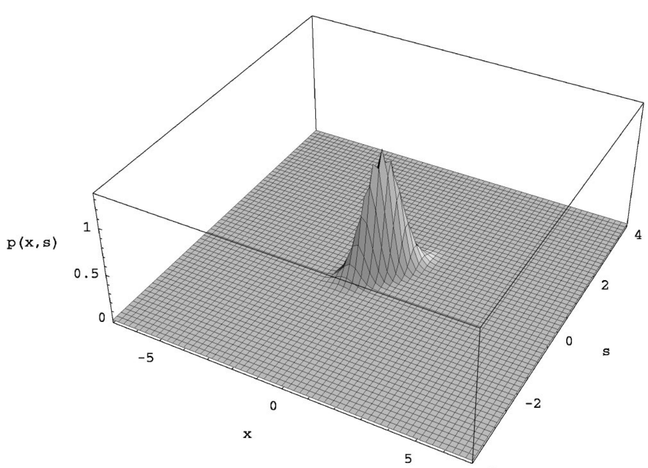

FIG. 10. Structure function $p(x,s)$ for the case $\beta=-\frac{1}{2}$ and $\gamma=2$.

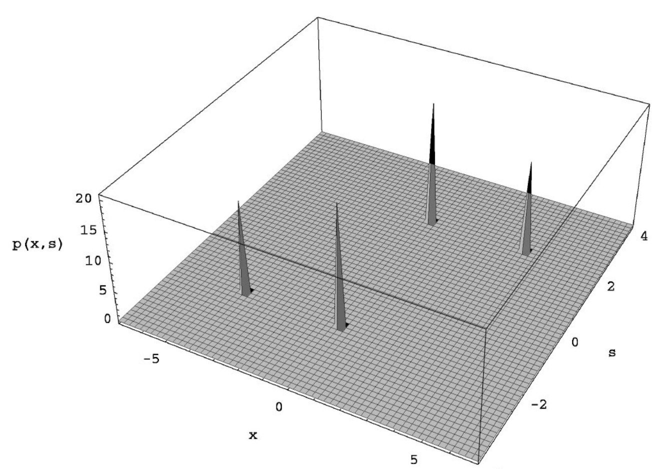

FIG. 12. Structure function $p(x,s,\beta,\gamma)$ for the case $\beta=16$ and $\gamma=2$.

$$\cos^N(k\gamma)=\begin{cases}\frac{1}{2^N}\binom{N}{\frac{N}{2}}+\frac{1}{2^{N-1}}\sum_{m=0}^{N/2-1}\binom{N}{m}\cos[S_m k\gamma], & N \text{ even}\\ \frac{1}{2^{N-1}}\sum_{m=0}^{N-1/2}\binom{N}{m}\cos[S_m k\gamma], & N \text{ odd,}\end{cases} \tag{4.4}$$

where we have defined $S_m=N-2m$. Hence, we have for $\tilde{Q}^N(k)$,

$$\tilde{Q}^N(k)=e^{-Nk^2/16\beta\gamma^2}\begin{cases}\frac{1}{2^N}\binom{N}{\frac{N}{2}}+\frac{1}{2^{N-1}}\sum_{m=0}^{N/2-1}\binom{N}{m}\cos[S_m k\gamma], & N \text{ even}\\ \frac{1}{2^{N-1}}\sum_{m=0}^{N-1/2}\binom{N}{m}\cos[S_m k\gamma], & N \text{ odd.}\end{cases} \tag{4.5}$$

We now recall a useful integral

$$\frac{1}{2\pi}\int_{-\infty}^{\infty}e^{-ikx}e^{-\alpha k^2}\cos(\beta k)dk=\frac{1}{\sqrt{4\alpha\pi}}\left[e^{-x^2+\beta^2/4\alpha}\cosh\left(\frac{\beta|x|}{2\alpha}\right)\right]. \tag{4.6}$$

Hence, calculating via Eqs. (2.5) and (4.5), we obtain

$$P_N(x)=\frac{1}{2^N}\sqrt{\frac{4\beta\gamma^2}{N\pi}}e^{-4\beta\gamma^2x^2/N}\begin{cases}\binom{N}{\frac{N}{2}}+2\sum_{m=0}^{N/2-1}\binom{N}{m}e^{-4\beta S_m^2\gamma^4/N}\cosh\left[\frac{8S_m\gamma^3|x|}{N}\right] & N \text{ even}\\ 2\sum_{m=0}^{N-1/2}\binom{N}{m}e^{-4\beta S_m^2\gamma^4/N}\cosh\left[\frac{8S_m\gamma^3|x|}{N}\right], & N \text{ odd.}\end{cases} \tag{4.7}$$

The expression in Eq. (4.7) is a completely analytic solution, which provides insight into the asymptotic behavior of $P_N(x)$.

Let us first consider the limit as $N\to\infty$. We immediately note that the prefactor in Eq. (4.7) is exactly the Gaussian form we expect, except for the factor of $1/2^N$. We therefore wish to show that the term in the brackets—regardless of whether N is even or odd—approaches 2^N in the large N limit. We consider the exponential weighting of the combinatorial factors

$$P_0=e^{-4\beta S_m^2\gamma^4/N}. \tag{4.8}$$

This is independent of position—the only dependence is on the parameter m. For a fixed *numerical* value of S_m (the value of m required will increase with N), the contribution of each term becomes *less* exponentially suppressed. In other words, as N becomes large, contributions of distant peaks become more and more important at every location in the space. Similarly, regardless of the value of x, for fixed *numerical* values of S_m, the argument of the cosh function will tend towards zero. Therefore for large N, both the exponential and cosh terms become arbitrarily close to one, allowing us to rewrite Eq. (4.7) as

$$P_N(x)\to\frac{1}{2^N}\sqrt{\frac{4\beta\gamma^2}{N\pi}}e^{-4\beta\gamma^2x^2/N}\times\begin{cases}\binom{N}{\frac{N}{2}}+2\sum_{m=0}^{N/2-1}\binom{N}{m}, & N \text{ even}\\ 2\sum_{m=0}^{N-1/2}\binom{N}{m}, & N \text{ odd.}\end{cases} \tag{4.9}$$

Noting that binomial coefficients are symmetric under the exchange of m and $N-m$, we may rewrite Eq. (4.9) as

$$P_N(x)\sim\frac{1}{2^N}\sqrt{\frac{4\beta\gamma^2}{N\pi}}e^{-4\beta\gamma^2x^2/N}\sum_{m=0}^{N}\binom{N}{m}. \tag{4.10}$$

The sum is simply equal to 2^N, so we therefore have for large N,

$$P_N(x)\sim\sqrt{\frac{4\beta\gamma^2}{N\pi}}e^{-4\beta\gamma^2x^2/N}=P_G(x), \tag{4.11}$$

which is the Gaussian form, as expected from the central limit theorem.

It is worth mentioning that the general case of convergence to a Gaussian probability distribution for a symmetric random walk with a nonzero third moment for its jump probabilities was investigated a long time ago. The Berry-Esséen theorem [14] which applies to this case gives a criterion for the number of steps n required to achieve a desired level of convergence to the Gaussian and it states that

$$n \gtrsim \frac{25}{4} \frac{\langle |x|^3 \rangle^2}{\langle x^2 \rangle^3} \frac{1}{\epsilon^2}. \tag{4.12}$$

Recently, Montegna and Stanley [15] investigated the issue of slow convergence to Gaussian behavior of the truncated Lévy flight process, where $p(x) \approx 1/x^{1+\alpha}$ for $|x| < L$, $\alpha < 2$ and $p(x) = 0$ for $|x| > L$. They found that convergence is achieved typically after $n \approx 10^4$ steps as compared to $n \approx 10$ for common distributions. Schlesinger [16] showed that this number scales as $n \sim L^{\alpha}$.

We have applied this criterion to our double-well step size distribution and obtained the following general formula:

$$n \gtrsim \frac{25}{4} \frac{1}{\epsilon^2} \frac{\Gamma\left(\frac{1}{2}\right)\left[\Gamma\left(\frac{1}{2}\right)\right]^2}{\Gamma\left(\frac{3}{2}\right)^3} \frac{D_{-1/2}(\xi) D_{-2}{}^2(\xi)}{D_{-3/2}{}^3(\xi)}, \tag{4.13}$$

where $\xi = \alpha/\sqrt{2\beta}$. In order to find a numerical estimate of the minimum number of steps required we used two separate limiting cases depending on the form of the single-step probability distribution. For α small and β large ($\xi \ll 1$), i.e., a strongly non-Gaussian case, we use [17]

$$D_{-a-1/2}(\xi) \sim \left[\frac{\sqrt{\pi} 2^{-a/2-1/4}}{\Gamma\left(\frac{3}{4} + \frac{a}{2}\right)} \right] \exp(\mp \sqrt{a} \xi), \tag{4.14}$$

while in the opposite case, i.e., α large and β small ($\xi \gg 1$) we approximate the parabolic cylindrical function by

$$D_{-a-1/2}(\xi) \sim \exp\left(-\frac{\xi^2}{4} \right) \xi^{-a-1/2}. \tag{4.15}$$

Consequently, the Berry-Esséem criterion gives a numerical prediction in terms of the number of steps required at a given confidence level ϵ. First of all, it is readily seen that in both cases, to the first order of approximation, there is no scaling with respect to α or β. We find that

$$n \sim \frac{1}{\epsilon^2} \begin{cases} \frac{150}{\pi^2} \times \frac{\Gamma^3(1.25)}{\Gamma(-0.75)} \simeq 12 & \text{for } \xi \ll 1 \\ \frac{50}{\pi} \simeq 16 & \text{for } \xi \gg 1. \end{cases} \tag{4.16}$$

Expecting, for example, that the difference between the double well and the Gaussian step distribution to be less than $\epsilon \simeq 0.25$, we present the required number of steps to range between 192 and 256 depending on whether we deal with strongly or weakly quartic cases. Below, we have attempted to verify this prediction directly by assessing the separation between neighboring peaks in the cummulative probability distribution.

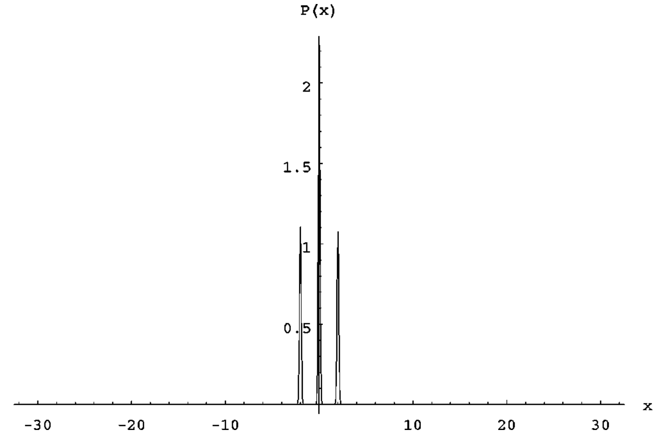

FIG. 13. Probability distribution $P(x)$ for finding a random walker at location x after $N = 2$ steps ($\beta = 8$, $\gamma = 1$).

The exponential form P_0 in Eq. (4.8) is also useful in this regard since it allows us to derive an expression for a "critical value" of N, at which adjacent peaks begin to contribute significantly to each others amplitudes. Without loss of generality, consider the case of the peak at the origin for N even. Singling out the contributions of the adjacent peaks by setting $S_m = 2$, we solve for the value of N at which the exponent is $-1/\sqrt{2}$. At this point the exponential will be $\approx \frac{1}{2}$, such that the combined contribution of the two neighboring peaks is of comparable magnitude to that of the peak at the origin itself. We obtain the following expression for N_{crit}:

$$N_{crit} = 16 \sqrt{2} \beta \gamma^4. \tag{4.17}$$

This critical value for N can be expressed in two other useful forms. In terms of the original parameters α and β of Eq. (2.16), we can write

$$N_{crit} = \frac{4\sqrt{2}\alpha^2}{\beta}, \tag{4.18}$$

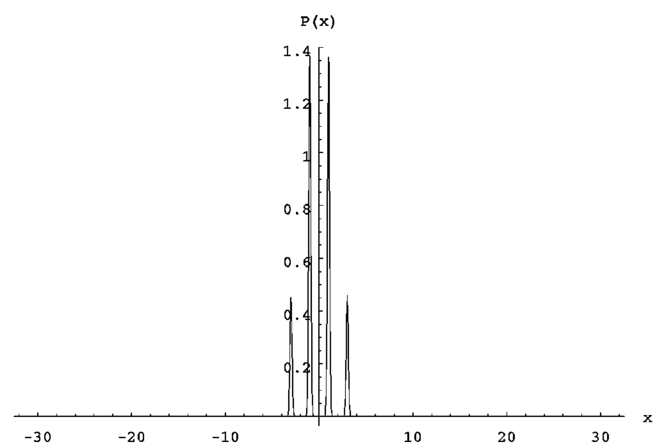

FIG. 14. Probability distribution $P(x)$ for finding a random walker at location x after $N = 3$ steps ($\beta = 8$, $\gamma = 1$).

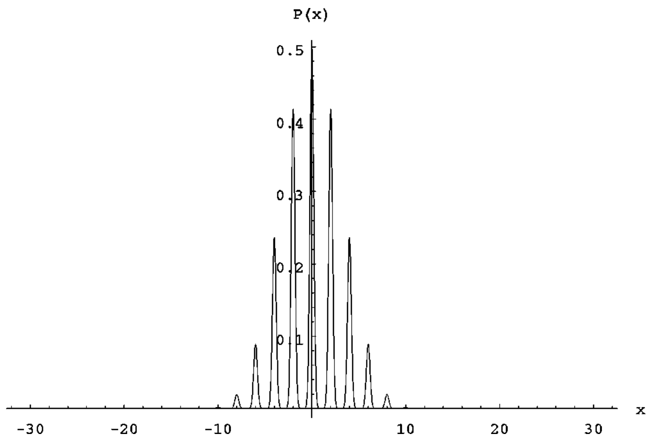

FIG. 15. Probability distribution $P(x)$ for finding a random walker at location x after $N=10$ steps ($\beta=8,\ \gamma=1$).

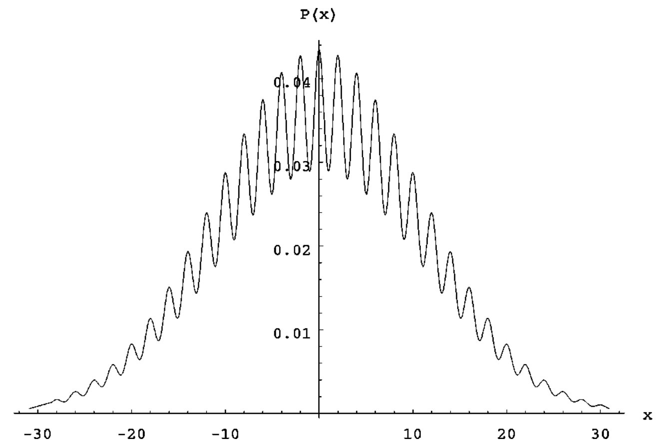

FIG. 17. Probability distribution $P(x)$ for finding a random walker at location x after $N=120$ steps ($\beta=8,\ \gamma=1$).

while in terms of the width σ and the separation Δ of the peaks, we may rewrite Eq. (4.17) as

$$N_{crit}=\frac{\Delta^2}{\sqrt{2}\sigma^2}. \tag{4.19}$$

This critical value should correspond roughly to the number of steps which our walker must take before the probability distribution becomes relatively close to a Gaussian form. We shall examine this hypothesis below.

As a typical example, we consider a set of plots of $P(x)$ for the case $\beta=8,\ \gamma=1$. By Eq. (4.17), we expect $P(x)$ to be approximately Gaussian after about 181 steps. This convergence is explored graphically in Figs. 13–19 below. This lies within the range of our estimated number of steps and the case choosen here as an example corresponds to $\xi=4$.

V. CONCLUSIONS

In this paper, we have examined the issue of the asymmptotic convergence of a random walk process with a double-peak distributed step size to its Gaussian limit. We used a combination of analytical, approximate, and numerical methods.

Any attempt to obtain an exact analytic result for $P(x)$ requires a simple closed form solution of $Q(k)$. Finding such a solution is difficult (see Appendix B), although we know it must be even in k, and a function (or combination of functions) of the form ${}_1F_1(\alpha,\beta,f(k))$. At present, the only exact analytic solution we have is an infinite series of hypergeometric functions. Without a simple form of an exact solution of $Q(k)$, attempts to examine $P(x)$ numerically have proven very challenging.

However, we do have an approximate form (4.1) of $w(s)$, consisting of two narrowly peaked Gaussians at $\pm\gamma$, which is valid over a wide range of parameters. Through this approximation, we are able to study the asymptotic behavior of the system. We are able to see both how we retrieve the binomial distribution in the limit $\beta\to\infty$, and how we obtain a Gaussian form in the large N limit. We have shown how our asymptotic results are consistant with the predictions made in the general case by the Berry-Esséen theorem.

The ideal extension of the work in this paper would be to find a simple exact form for $Q(k)$ to allow one to solve exactly for $P(x)$. However, in the eventuality that such a

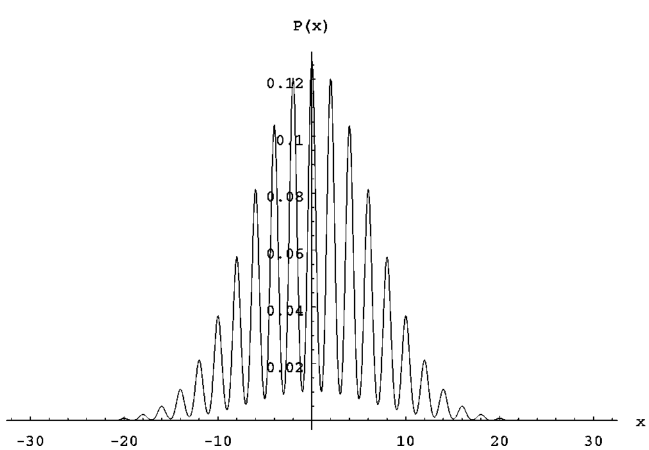

FIG. 16. Probability distribution $P(x)$ for finding a random walker at location x after $N=40$ steps ($\beta=8,\ \gamma=1$).

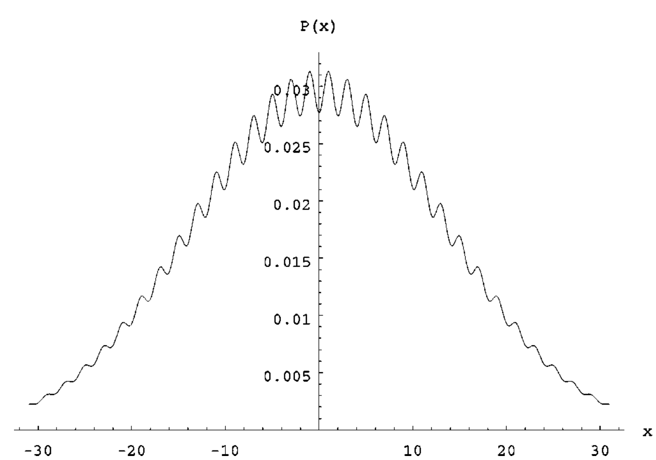

FIG. 18. Probability distribution $P(x)$ for finding a random walker at location x after $N=181$ steps ($\beta=8,\ \gamma=1$).

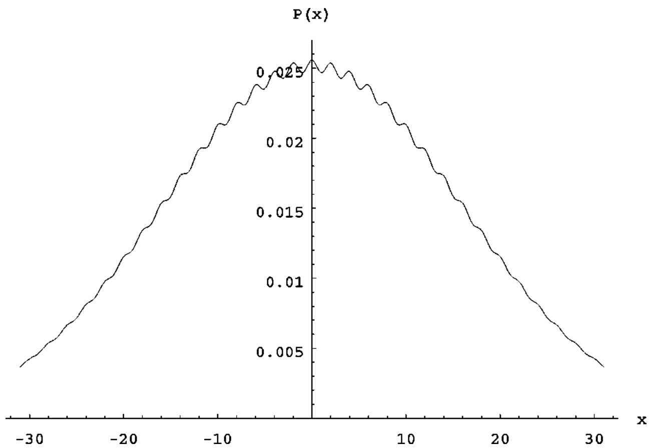

FIG. 19. Probability distribution $P(x)$ for finding a random walker at location x after $N=250$ steps ($\beta=8,\ \gamma=1$).

solution is not possible, further progress can be made with the approximation of Eq. (4.1). One line of attack would be to improve the approximation to take into account the asymmetry of the peaks. The asymmetry appears to be of a similar form to the first derivative of a Gaussian; however, at the time of writing, no particular improvement on the approximation given in Eq. (4.1) has been found. Another possibility would be to find another approximation for the range where $\beta<0$, which would hopefully cover another wide swath of the parameter space, allowing for study of the asymptotic behavior of the single peak case.

It has recently emerged that a practical application of this type of process might be to the protein folding problem or more directly to the problem of calculating the radius of gyration of a peptide composed of identical monomers. The case of a random chain is often analyzed using the Gaussian radial probability distribution in three-dimensional space is [18]

$$P(r)=4\pi r^2(2\pi\sigma^2)^{-3/2}\exp\left(\frac{-r^2}{2\sigma^2}\right), \qquad (5.1)$$

where $\sigma^2=Na^2/3$. Here, N is the number of monomer units within the chain, a the unit length. Then, using well-known formulas in Gaussian statistics one finds that the average end-to-end distance $\langle r_{\rm ee}\rangle$ scales as

$$\langle r_{\rm ee}\rangle=\sqrt{\frac{8}{3\pi}}aN^{1/2}, \qquad (5.2)$$

while the variance of this variable is given by

$$\langle r_{\rm ee}^2\rangle=Na^2. \qquad (5.3)$$

However, in general, the mean square end-to-end distance $\langle r_{\rm ee}^2\rangle$ is known to satisfy the more complicated relationship

$$\langle r_{\rm ee}^2\rangle=AN^{\nu}, \qquad (5.4)$$

where $1\leqslant\nu\leqslant2$ depending of the case [19,20]. Using the non-Gaussian probability distribution instead of Eq. (5.1) leads to a range of exponents with the Gaussian value being only one limiting case [21]. Furthermore, this assumption is fully justified by the local energy dependance on the rotation angle for the common peptide nitrogen–carbon and carbon-carbon bonds (see Fig. 20). Finally, recent advances in the protein folding problem indicate that the potential landscape possesses a heirachy of barrier heights and, so far, has been emulated by the so-called spin-glass model [22,23]. We show

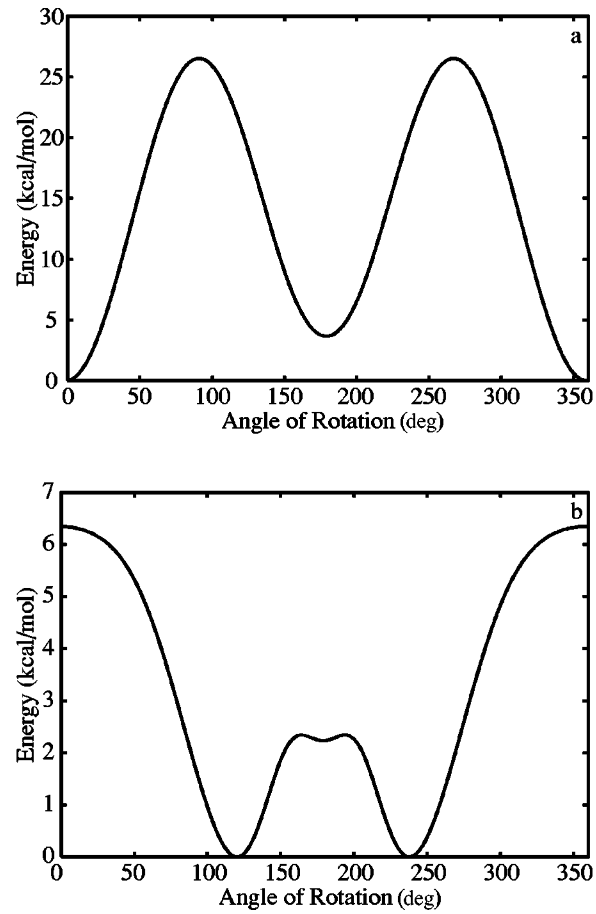

FIG. 20. The approximate energy changes associated with rotation about the (a) N–C (ϕ) and (b) C_α–C (ψ) bonds in a G–G dipeptide.

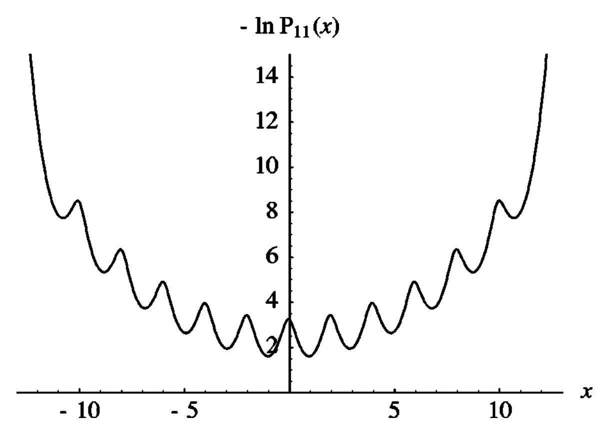

FIG. 21. A plot of $-\ln P(x)$ after 11 steps for $w(s)$ given by Eq. (2.16) when $\alpha=-16$ and $\beta=8$ ($\gamma=1$).

in Fig. 21 that the effective potential stemming from the non-Gaussian distributed random walks shows some resemblence to it. The actual protein energy landscapes show a self-similar hierarchical structure that is absent from Fig. 21. However, one can envisage random walk process on more than a single length scale with large, medium, and small structural elements of a protein making choices about their orientation in space. We intend to purse this analogy further in a future publication.

ACKNOWLEDGMENTS

R. Teshima was very helpful both for his suggestions of possible numerical techniques and for his extensive knowledge of special functions. E. Finley is recognized for his assistance in ruling out a couple of possible hypergeometric forms for $Q(k)$.

APPENDIX A: INTEGRALS LEADING TO PARABOLIC CYLINDER FUNCTIONS

Following general results published for non-Gaussian models of thermodynamic fluctuations [24,25] and using the integral

$$\int_0^\infty x^{2k\nu-1}\exp(-ax^{4k}-bx^{2k})dx$$

$$=(2k)^{-1}(2a)^{-\nu/2}\Gamma(\nu)\exp\left(\frac{b^2}{2a}\right)D_{-\nu}\left(\frac{b}{\sqrt{2a}}\right), \quad \text{(A1)}$$

where $D_{-\nu}(x)$ is the parabolic cylinder function [17], we find that the probability distribution given by

$$P(x)=C\exp[-\lambda_2(x-\overline{x})^2-\lambda_4(x-\overline{x})^4] \quad \text{(A2)}$$

is characterized by the normalization factor

$$C^{-1}=(2\lambda_4)^{-1/4}\Gamma\left(\frac{1}{2}\right)\exp\left(\frac{\lambda_2^2}{8\lambda_4}\right)D_{-1/2}\left(\frac{\lambda_2}{\sqrt{2\lambda_4}}\right) \quad \text{(A3)}$$

and its nth moment M_n,

$$M_n\equiv C^{-1}\int_{-\infty}^{\infty}(x-\overline{x})^nP(x)dx$$

$$=(2\lambda_4)^{-n/2}\frac{\Gamma\left(\frac{n+1}{2}\right)}{\Gamma\left(\frac{1}{2}\right)}\frac{D_{-n+1/2}\left(\frac{\lambda_2}{\sqrt{2\lambda_4}}\right)}{D_{-1/2}\left(\frac{\lambda_2}{\sqrt{2\lambda_4}}\right)}. \quad \text{(A4)}$$

APPENDIX B: HYPERGEOMETRIC FUNCTIONS

As we have seen in Sec. II, our form for $Q(k)$ is an infinite series of hypergeometric functions. Let us, therefore, recall some of the properties of functions of this type. A generalized hypergeometric function of order (n,d) is defined as

$$ {}_nF_d(\alpha_1,\ldots,\alpha_n,\beta_1,\ldots,\beta_d,x)=\sum_{m=0}^{\infty}\frac{(\alpha_1)_m\cdots(\alpha_n)_m}{(\beta_1)_m\cdots(\beta_d)_m}\frac{x^m}{m!}, \quad \text{(B1)}$$

where the Pochammer symbol $(\alpha)_m$ is defined as

$$(\alpha)_m=\prod_{j=0}^{m-1}(\alpha+j). \quad \text{(B2)}$$

Almost all special functions of mathematical physics are merely special cases of hypergeometric functions. In particular, the exponential function relates as

$$e^{f(x)}={}_0F_0(f(x)). \quad \text{(B3)}$$

It is a general and useful fact that calculus operations on a hypergeometric function of order (n,d) increase the order to $(n+1,d+1)$. Therefore, we may expect our $Q(k)$ to be expressable in terms of functions of the form ${}_1F_1(\alpha,\beta,f(k))$. At present, we have identified 17 different choices $f(k)$ which do not solve the problem. However, a possible lead comes from one of the many addition theorems [13] which exist for hypergeometric functions of order (1,1), namely,

$${}_1F_1(\alpha,\beta,x+y)=\left(\frac{x}{x+y}\right)^{\alpha}\sum_{n=0}^{\infty}\frac{(\alpha)_ny^n}{n!(x+y)^n}{}_1F_1(\alpha+n,\beta,x). \quad \text{(B4)}$$

Our solution for $Q(k)$ given in Eq. (2.21) above can be rewritten as a sum of four terms of the form (neglecting multiplicative factors)

$$\sum_{n=0}^{\infty}\frac{k^{4n}}{4n!}\beta^{-n}\Gamma\left(\frac{1}{4}+n\right){}_1F_1\left(\frac{1}{4}+n,\frac{1}{2},\beta\gamma^4\right), \quad \text{(B5)}$$

which differs only in the factor of $(4n)!$ in the denominator. This suggests that we can expect a function $f(k)$ of a form similar to

$$f(k)\approx\frac{\beta^2\gamma^4}{\beta-k^4} \quad \text{(B6)}$$

in a representation of $Q(k)$ in terms of a finite number of hypergeometric functions.

[1] S. Chandresekhar, Rev. Mod. Phys. **15**, 1 (1943).
[2] F. Reif, *Statistical and Thermal Physics* (McGraw-Hill, Toronto, Canada, 1965).
[3] A. Bunde and S. Havlin, *Fractals and Disordered Systems* (Springer-Verlag, Berlin, 1996).
[4] A. I. Khinchin, *Mathematical Foundations of Statistical Mechanics* (Dover, New York, 1949).
[5] G.A. Baker, Jr., Phys. Rev. Lett. **60**, 1844 (1988).

[6] M.S.S. Challee and J.H. Hetherington, Phys. Rev. Lett. **60**, 77 (1988).
[7] M. Gitterman, Physica A **177**, 324 (1991).
[8] V. de Giorgio and M.O. Scully, Phys. Rev. A **2**, 1170 (1970).
[9] R.D. Astumian, Science **276**, 917 (1997).
[10] S. Leibler and D.A. Huse, J. Cell Biol. **21**, 1357 (1993).
[11] G. Nicolis and I. Prigogine, *Exploring Complexity* (W. H. Freeman, New York, 1989).
[12] I. S. Gradshteyn and I. M. Ryzhik, *Table of Integrals, Series, and Products*, 5th ed. (Academic Press, Toronto, Canada, 1994).
[13] L. J. Slater, *Confluent Hypergeometric Functions* (Cambridge University Press, Cambridge, 1960).
[14] W. Feller, *An Introduction to Probability Theory and Its Applications* (Wiley, New York, 1966), Vol. 2, Chap. XXI 8, p. 525.
[15] R.N. Montegna and H.E. Stanley, Phys. Rev. Lett. **73**, 2946 (1994).
[16] M.F. Schlesinger, Phys. Rev. Lett. **74**, 4959 (1995).
[17] *Handbook of Mathematical Functions*, edited by M. Abramowitz and I. A. Stegun (Dover, New York, 1965).
[18] D. Boal, *Mechanics of the Cell* (Cambridge University Press, Cambridge, 2001).
[19] M. Doi and S. F. Edwards, *The Theory of Polymer Dynamics* (Oxford University Press, Oxford, 1986).
[20] H. Yamakawa, *Modern Theory of Polymer Solutions* (Harper and Row, New York, 1971).
[21] J. M. Dixon, J. A. Tuszynski, and P. A. Clarkson, *From Nonlinearity to Coherence* (Oxford University Press, Oxford, 1997).
[22] K. Binder and A.P. Young, Rev. Mod. Phys. **58**, 801 (1986).
[23] M. Kursynski, Prog. Biophys. Mol. Biol. **69**, 23 (1998).
[24] J.A. Tuszynski, M.J. Clouter, and H. Kieffe, Phys. Lett. **108A**, 272 (1985).
[25] J.A. Tuszynski, M.J. Clouter, and H. Kieffe, Phys. Rev. B **33**, 3423 (1986).