Study of the shape of random walks

1994 J. Phys. A: Math. Gen. 27 7015

(http://iopscience.iop.org/0305-4470/27/21/017)

J. Phys. A: Math. Gen. 27 (1994) 7015–7034. Printed in the UK

Study of the shape of random walks

S J Sciutto

Laboratorio de Física Teórica, Departamento de Física, Universidad Nacional de La Plata, C C 67–1900 La Plata, Argentina

Received 31 March 1994

Abstract. A comprehensive study of the shape of discrete random walks, considering the general case of arbitrary length and any space dimension, is presented. The probability distributions for several magnitudes such as the principal inertia moments, the asphericity and the angle between the principal axis of inertia and the end-to-end vector are evaluated numerically. In most cases, and especially in low-dimensional spaces, these probability distributions spread widely and possess a non-zero skewness. This implies that any description of the shapes of random walks which is based only on mean values of related magnitudes is incomplete. This situation does not vary when random walks of large lengths are considered since the relative fluctuations (ratio between the standard deviation and the mean value) do not go to zero in that limit. On the other hand, when the space dimension grows, the distributions become more peaked, reaching the expected asymptotic limit for infinite dimension. The present study also indicates that the probability distributions for the principal inertia moments have approximately the form of chi-squared distributions. Analysing the probability distribution for the asphericity, it is shown that the distributions for the different moments are not completely independent.

1. Introduction

Random walks and related objects are of interest in many branches of physics such as the theory of polymer molecules [1] or quantum field theory [2]. In particular, the problem of determining the characteristics of the shapes of random walks has been of interest to scientists for many years and motivated a number of works [3–9]. In the pioneering papers of [3] an analysis was made of the mean values of the principal inertia moments for three-dimensional random walks, obtaining the result that they are not in the proportion 1:1:1, immediately suggesting a non-spherical shape for these objects. This result triggered a series of analyses [3–9] trying to determine the origin, characteristics and implications of that property of random walks. Particularly, the interest in giving a more precise quantitative measure of the shapes of these objects motivated the definition of other magnitudes. An important example being the asphericity, introduced some time ago [10]. This scalar quantity takes values in the interval [0, 1], 0 for a perfectly spherical object and 1 for a rod-shaped one. In [10] it was explicitly shown that a quantity similar to the mean value of the asphericity is not zero for unrestricted random walks in spaces of arbitrary dimension, a result that leads once more to the conclusion that random walks do not possess a spherically symmetric shape.

In [5] it was shown that in an infinite-dimensional space, all open unrestricted random walks possess the same shape. In the case of large but finite dimension, the shapes (quantitatively measured by some quantity like the asphericity, for example) will distribute

0305-4470/94/217015+20$19.50

around a most probable shape. The distribution will present a peak around the most probable configuration whose sharpness decreases when the dimension is lowered.

In a more recent work [7], qualitatively similar results were obtained when studying other kind of random walks—like self-avoiding ones, for example—and/or other magnitudes, i.e. the angle between one of the principal inertia axes and the end-to-end vector.

The method used in most of the previously mentioned works consists of the evaluation of *mean values* of certain magnitudes (inertia moments, asphericity, etc). In general, the study of other statistical parameters like the deviations around the mean, the most probable value, etc, is omitted.

In cases like, for example, thermodynamical systems, mean values can describe adequately the behaviour of the corresponding magnitude because the relative fluctuations (ratio between standard deviation and mean value) vanish in the limit of infinite system size. However, it has not been shown that something similar to such a limit in the problem of determining the shapes of random walks exists.

It would, therefore, be important to perform a complete study of the different statistical parameters in order to acquire a detailed description of the shape (or distribution of shapes) of a set of random walks.

The main scope of this work is to present a comprehensive study of the shape of random walks. We make the analysis of unrestricted discrete random walks of arbitrary length in spaces of different dimensions. Our study is mainly focused on the probability distributions of different magnitudes, namely, the principal inertia moments, the asphericity and the angle between the principal axis of inertia and the end-to-end vector. It should be mentioned, however, that in the case of Gaussian (not discrete) random walks, there exist previous works which study the probability distributions of related quantities. For example, reference [4] contains an analytical study of the probability distributions of the principal inertia moments of two-dimensional rings; in reference [8] there is a detailed study of probability distributions for the asphericity of unrestricted and self-avoiding random walks in three dimensions, and finally reference [11] reviews the theory of the distribution function of the radius of gyration (see below) of Gaussian random walks in a three-dimensional space. In this last work [11] no explicit mention is made of the non-spherical shape of the random walks; and none of the works [4, 8, 11] consider explicitly the case of discrete random walks.

Among other results, we find that all the magnitudes considered present broad distributions with significative fluctuations and non-zero skewness, especially for low spatial dimension (two or three). When the dimension is increased, the distributions become sharply peaked around a single most probable value, in concordance with reference [6].

Even if some of the cited works study the ratios between different inertia moments, we are not going to consider such quantities here because they are not practical when considering spaces of more than three dimensions and the results obtained from analysing them are not likely to be qualitatively different from the ones presented in this work.

Our paper is organized as follows. In section 2 we define the different magnitudes used in the calculations and discuss some analytical results for the limit of infinite dimension. In section 3 we describe the numerical method used and present the results obtained, including the analysis of the probability distributions, the comparison with analytical models, and the study of mean values and dispersions. Finally, in section 4, we state our conclusions and final remarks.

2. Describing the shape of random walks

Let us consider a random walk of s steps in a d-dimensional space. If $r_0 = 0$ is the initial position and r_α, $\alpha = 1, \ldots, s$ ($r_\alpha \equiv (x_{1\alpha}, x_{2\alpha}, \ldots, x_{d\alpha})$), represent the positions within the walk, we can define the random walk stating that

$$r_\alpha = r_{\alpha-1} + \epsilon_\alpha \tag{1}$$

where ϵ_α, $\alpha = 1, \ldots, s$, is a random step vector. We shall consider for simplicity discrete unrestricted random walks with coordination number $2d$. Therefore, if $B = \{e_1, e_2, \ldots, e_d\}$ is an orthonormal basis for the d-dimensional space, ϵ_α may be any of the vectors $\pm e_1, \pm e_2, \ldots, \pm e_d$, each one with probability $(2d)^{-1}$.

For such a random walk we can define the *centre of mass*

$$r_{CM} = \frac{1}{s+1} \sum_{\alpha=1}^{s} r_\alpha \tag{2}$$

and the *inertia matrix*

$$T_{ij} = \frac{1}{s+1} \sum_{\alpha=0}^{s} (x_{i\alpha} - x_{CMi})(x_{j\alpha} - x_{CMj}) \qquad 1 \leqslant i \leqslant d \quad 1 \leqslant j \leqslant d\,. \tag{3}$$

This matrix is symmetric and positive definite. It possesses d positive eigenvalues $\lambda_1 \geqslant \lambda_2 \geqslant \cdots \geqslant \lambda_d$, and an orthogonal set of d eigenvectors u_k, $k = 1, \ldots, d$ such that $T u_k = \lambda_k u_k$.

The trace of the inertia matrix is a well known magnitude called the *radius of gyration*

$$S^2 = \sum_{k=1}^{d} T_{kk} = \sum_{k=1}^{d} \lambda_k \tag{4}$$

which gives a quantitative measure of 'how large' the random walk is. For fixed s, S^2 takes values from (almost) zero to a maximum S^2_{max}. The upper bound for S^2 can be evaluated in the following way: from (2)–(4) it follows that

$$S^2 = \frac{1}{s+1} \sum_{\alpha=1}^{s} r_\alpha^2 - r_{CM}^2\,. \tag{5}$$

From (1) it follows that for $\alpha \geqslant 1$

$$r_\alpha = \sum_{\beta=1}^{\alpha} \epsilon_\alpha \tag{6}$$

where all the vectors ϵ_α, $\alpha = 1, \ldots, s$, are unitary. Replacing into (2) and operating conveniently, the centre of mass can be expressed in terms of the vectors ϵ_α:

$$r_{CM} = \sum_{\alpha=1}^{s} \left(1 - \frac{\alpha}{s+1}\right) \epsilon_\alpha\,. \tag{7}$$

Using (6) and (7) to replace r_α and r_{CM} in (5), and performing some algebraic manipulations, one obtains

$$S^2 = \tfrac{1}{2}s - \frac{s(2s+1)}{6(s+1)} + \Sigma_2 \tag{8}$$

with

$$\Sigma_2 = \frac{2}{s+1} \sum_{\alpha=2}^{s} \left(1 - \frac{\alpha}{s+1}\right) \sum_{\beta=1}^{\alpha-1} \beta\, \epsilon_\alpha \cdot \epsilon_\beta \tag{9}$$

for all $s \geqslant 2$; for $s < 2$ $\Sigma_2 = 0$. The scalar products $\epsilon_\alpha \cdot \epsilon_\beta$ are less than or equal to one for all α and β since the step vectors ϵ_α are unitary. Therefore, it is evident that

$$\Sigma_2 \leqslant \frac{2}{s+1}\sum_{\alpha=2}^{s}\left(1-\frac{\alpha}{s+1}\right)\sum_{\beta=1}^{\alpha-1}\beta = \frac{s(s-4)}{12}+\frac{s(2s+1)}{6(s+1)}. \tag{10}$$

From (8) and (10) it follows that the radius of gyration possesses the following upper bound:

$$S^2 \leqslant S^2_{\max} = \tfrac{1}{12}s(s+2). \tag{11}$$

Equations (8)–(11) are valid for all space dimensions, path lengths and coordination numbers (they are valid also in the continuum), provided that the step vectors are unitary. Notice also that $S^2_{\max}$ is equal to the largest inertia moment of a rod of length s.

For an object with spherical shape the inertia eigenvalues are equal. In order to show that the random walks belonging to a given set possess a non-spherical shape, one must establish that these eigenvalues are somehow different. Many derived magnitudes can be defined which are useful to this purpose. A relevant example of such a magnitude is the *asphericity* defined by [10]:

$$A = \frac{\sum_{i=1}^{d-1}\sum_{j=i+1}^{d}(\lambda_i-\lambda_j)^2}{(d-1)\left[\sum_{i=1}^{d}\lambda_i\right]^2}. \tag{12}$$

It possesses the following properties: (i) $0 \leqslant A \leqslant 1$, (ii) $A = 0$ for spherical shape, and (iii) $A = 1$ for a rod-shaped object.

It is worthwhile mentioning that the present definition for the asphericity differs slightly from the usual one [10], where A is defined as a ratio of mean values. The reason to take our alternative definition is that it permits the calculation of a probability distribution for the asphericities of a given set of random walks which gives a much more complete information than a single scalar parameter.

We want also to mention that it is not difficult to demonstrate that the asphericity can be put in the following equivalent form:

$$A = \frac{1}{d-1}\left[d\left(\sum_{i=1}^{d}\lambda_i^2\right)\left(\sum_{i=1}^{d}\lambda_i\right)^{-2}-1\right] \tag{13}$$

which is particularly adequate for numerical calculations.

It is also interesting to study the behaviour of the first quadrant angle, Θ, between the path ends and the inertia axis corresponding to the largest eigenvalue λ_1 [7]:

$$\cos\Theta = \frac{|\boldsymbol{r}_s \cdot \boldsymbol{u}_1|}{\|\boldsymbol{r}_s\|\,\|\boldsymbol{u}_1\|}. \tag{14}$$

The angle Θ will present an uniform distribution in the interval $\left[0, \frac{1}{2}\pi\right]$ if there are no correlations between $\boldsymbol{r}_s$ and $\boldsymbol{u}_1$.

Notice that when we refer to some property of random walks, such as the 'shape', we are speaking about a statistical analysis over a population of random walks. There are $K = (2d)^s$ possible random walks of length s in a d-dimensional space. If the walks are generated completely at random (like in a Monte Carlo simulation, for instance), each single random walk possesses a probability $1/K$ of being generated. Given a certain magnitude X defined for each random walk (such as the principal inertia moments, the asphericity, etc), it is possible to define a probability distribution for X, $P_X(x)$ with x real, such that $P_X(x)\,\mathrm{d}x$ gives the probability that $x \leqslant X \leqslant x + \mathrm{d}x$. To evaluate $P_X(x)\,\mathrm{d}x$ one just needs to count the number of random walks which verify the condition $x \leqslant X_{\mathrm{RW}} \leqslant x + \mathrm{d}x$ and divide it by K.

In a Monte Carlo simulation, a large number N of independent random walks is generated. N should be large enough to allow the set of random walks obtained be a representative sample of the entire population. Then a probability distribution can be *estimated* building a frequency histogram for the corresponding magnitude. We generally will not distinguish explicitly between a theoretical probability distribution and its statistical estimation.

2.1. The $d \to \infty$ case

There is a particular example which clearly illustrates that random walks are not spherical and that there are correlations between r_s and u_1. It is the case of very large dimensions: $d \to \infty$, $d \gg s$.

Let us divide the random walks into two subsets: (i) all random walks such that the vectors ϵ_α, $\alpha = 1, \ldots, s$, are linearly independent, and (ii) the remaining random walks. Let n_i be the number of elements in subset i ($i = 1, 2$). We immediately see that $n_1 + n_2 = K$. Let p_i be the probability of generating a random walk belonging to subset i, then $p_i = n_i/K$, $i = 1, 2$.

It is possible to demonstrate [5] that in the limit $d \to \infty$, $n_2/n_1 \to 0$, $p_1 \to 1$, and $p_2 \to 0$. In other words, in this limiting case all random walks that appear with finite probability (i.e. in a Monte Carlo simulation) belong to subset 1.

It is also possible to demonstrate [5] that all the random walks belonging to subset 1 can be mapped one onto any other by a combination of rotations and reflections. Since rotations and reflections keep unchanged distances and angles, it follows that all these random walks possess the same principal inertia moments, asphericity and angle Θ, among other magnitudes.

One of the random walks which belong to subset 1 is that whose steps vectors are $\epsilon_\alpha = e_\alpha$, $\alpha = 1, \ldots, s$. It is not difficult to evaluate and diagonalize the inertia matrix of such an object, and therefore of any element of subset 1. In the case $d \gg s \gg 1$ the corresponding eigenvalues and eigenvectors are given by [5]

$$\lambda_k = \begin{cases} \dfrac{1}{4(s+1)}\left[\sin\left(\dfrac{k\pi}{2(s+1)}\right)\right]^{-2} & 1 \leqslant k \leqslant s \\ 0 & s < k \leqslant d \end{cases} \tag{15}$$

and

$$u_k = \sqrt{\frac{2}{s+1}} \sum_{j=1}^{s} \sin\left(\frac{jk\pi}{s+1}\right) e_j \tag{16}$$

respectively.

For large s one can safely replace $\sin x = x$ in (15) to get

$$\lambda_k \cong \frac{s}{\pi^2 k^2} \qquad k = 1, 2, \ldots . \tag{17}$$

A direct calculation permits us to evaluate the asphericity in the present case. From (13) and (17)

$$A = \frac{1}{d-1}\left[d\left(\sum_{i=1}^{s} \frac{1}{k^4}\right)\left(\sum_{i=1}^{s} \frac{1}{k^2}\right)^{-2} - 1\right] \tag{18}$$

which for $d \to \infty$ and $s \gg 1$ reduces to

$$A = \frac{\zeta(4)}{\zeta^2(2)} = \frac{2}{5} \tag{19}$$

where $\zeta(x)$ is Riemann's zeta function.

Θ can also be evaluated straightforwardly,

$$\cos\Theta = \frac{|r_s \cdot u_1|}{\|r_s\|\,\|u_1\|} = \frac{\sqrt{2/s}\sum_{j=1}^{s}\sin(j\pi/s)}{\sqrt{s}}$$
$$= \frac{\sqrt{2}}{s}\frac{\sin^2(\pi/2)}{\sin(\pi/2s)}$$
$$= \frac{2\sqrt{2}}{\pi} \cong 0.9003 \tag{20}$$

that is,

$$\Theta \cong 0.4503 \text{ radians} = 25°\,48'. \tag{21}$$

Observe that since we are considering the limit $d \gg s \gg 1$, results (19)–(21) do not depend on d and/or s. It is also important to notice that an analytic calculation of the asphericity (defined as a ratio of mean values) was presented in [10], considering similar conditions for the random walks but without imposing any restriction on d. The result obtained there, $\langle A\rangle = 2(d+2)/(5d+4)$, reduces to the result of (19) for large d, as expected because in this limit there is only one possible value for A which is, of course, equal to the mean value $\langle A\rangle$.

2.2. *Large but finite d*

For large but finite d, p_2 is different from zero and random walks with different shapes can appear. Therefore the different magnitudes will not take a definite value. If d is large the corresponding probability distributions will be sharply peaked around a certain value, approximately equal to the value for the $d \to \infty$ case already considered.

In this case the probability distribution for the principal inertia moments takes the form of a chi-squared distribution [6]

$$P_k(\lambda) = W_k\,\lambda^{d/2-1}\exp\left(-\frac{d\,\lambda}{2\alpha_{0k}}\right) \qquad 1 \leqslant k \leqslant s \tag{22}$$

where

$$\alpha_{0k} = \frac{1}{4(s+1)}\left[\sin\left(\frac{k\pi}{2(s+1)}\right)\right]^{-2}. \tag{23}$$

Imposing the normalization condition

$$\int_0^\infty P_k(\lambda)\,\mathrm{d}\lambda = 1 \qquad 1 \leqslant k \leqslant s \tag{24}$$

the normalizing constant W_k can be written in terms of a Gamma function:

$$W_k = \frac{1}{\Gamma(d/2)}\left(\frac{d}{2\alpha_{0k}}\right)^{d/2}. \tag{25}$$

To evaluate the probability distribution of any magnitude $G(\lambda_1,\ldots,\lambda_d)$—a function of the eigenvalues λ_k, such as the asphericity, for example—one needs the combined probability distribution $Q(\lambda_1,\ldots,\lambda_r)$ ($r = \min(d,s)$) of the non-zero eigenvalues. If P_G represents the probability distribution for G, then

$$P_G(x) = \int_0^\infty \mathrm{d}\lambda_1 \int_0^\infty \mathrm{d}\lambda_2 \ldots \int_0^\infty \mathrm{d}\lambda_r\,\delta(x - G(\lambda_1,\ldots,\lambda_r,0,\ldots,0))\,Q(\lambda_1,\ldots,\lambda_r) \tag{26}$$

where $\delta(x)$ is Dirac's delta distribution.

It is usual to assume that the probabilities for the different λ_k's are independent. Under such a hypothesis, Q can be put in the following form:

$$Q(\lambda_1, \ldots, \lambda_r) = \prod_{i=1}^{r} P_k(\lambda_k) . \tag{27}$$

Even with this assumption it is not possible, in general, to evaluate the distribution (26) for quantities such as the asphericity (see [6]). In practical cases it has to be evaluated numerically.

The quantities introduced in this section, the associated probability distributions and their respective parameters, are useful to describe, with acceptable detail, the shape of random walks. The next sections are devoted to their analysis.

3. Numerical simulations and their results

The Monte Carlo simulations were performed under the following conditions: keeping d and s fixed, a large number N of statistically independent random walks with the characteristics of the already defined in section 2 were generated. For each random walk, the eigenvalues, $\lambda_1, \ldots, \lambda_d$, the eigenvectors, $u_1, \ldots, u_d$, the asphericity A and the angle Θ were evaluated. The N samples obtained in this process were used to perform several statistical analysis. In almost every calculation we have $N \geqslant 10^5$.

3.1. The inertia moments

Let us begin the study of the numerical data by considering the probability distribution for the principal inertia moments, λ_k. In figure 1 we can see a typical example; the case $d = 3$, $s = 100$, $k = 1$. The histogram plotted as a full curve represents the Monte Carlo data. The dotted curve represents a similar histogram corresponding to the probability distribution of (22). We were very careful when normalizing both histograms. In fact, we proceeded as follows:

(i) a certain number of frequency intervals $[x_i, x_{i+1})$, $i = 1, \ldots, m$, such that $[x_1, x_{m+1}]$ represents the region of interest of the respective variable (λ_1 in the present case), is defined.

(ii) The N Monte Carlo samples are used to evaluate the m frequencies f_i, $i = 1, \ldots, m$, which represent the number of times the variable happened to lie in the corresponding interval $[x_i, x_{i+1})$. These frequencies are normalized to make their sum equal to one: let h_i, $i = 1, \ldots, m$, be the normalized frequencies, then

$$h_i = \frac{f_i}{\sum_{j=1}^{m} f_j} . \tag{28}$$

(iii) A certain number N' ($N' = N$ in most of our simulations) of pseudo-random numbers distributed accordingly with the theoretical distribution (22) is generated, and processed similarly to the Monte Carlo data of the previous step, using the same frequency intervals defined at step (i). Frequencies f_i', $i = 1, \ldots, m$, and normalized frequencies

$$h_i' = \frac{f_i'}{\sum_{j=1}^{m} f_j'} \tag{29}$$

are evaluated as in step (ii).

(iv) The normalized frequencies h_i (full curves) and h_i' (dotted curves) are plotted in a single graph.

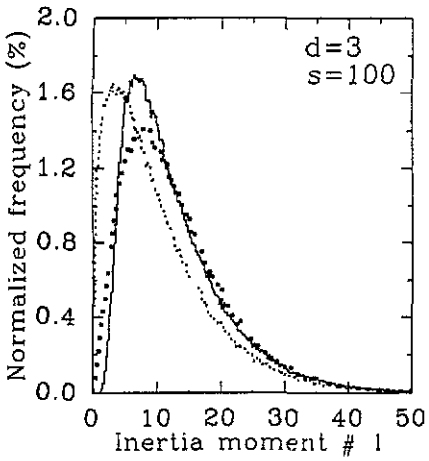

Figure 1. Frequency histograms for the largest inertia moment, λ_1, in the case $d = 3$, $s = 100$. The histograms with full curves correspond to Monte Carlo data. The histograms in dotted curves correspond to the distributions of (22) (small dots), and (31) (large dots). All the histograms are normalized accordingly with (28) and (29). In all cases $N = N' = 3\times10^5$.

As we can see in figure 1, the agreement between the two histograms is not complete. This is not surprising since (22) is obtained under the assumption $d \gg s \gg 1$ which certainly is not verified in the present example. It is worthwhile mentioning that the same case was analysed in [6], albeit evaluating a much smaller number of samples (10^4) than in our simulation. In that work both histograms seem to present a remarkably close agreement [6, figure 10]. It is unclear to us, however, how both curves were normalized, and whether or not both normalizations are compatible with a probability distribution normalization, i.e. equation (24).

We compared the Monte Carlo distributions for different eigenvalues λ_k and noticed that the probability distribution of (22) approximates well the simulation data just in those situations where d is sufficiently large. In general, the histogram corresponding to the distribution (22) differs from the Monte Carlo distribution and this also implies that there are differences between the mean values $\langle\lambda\rangle$ and standard deviations σ_λ corresponding to each distribution. We were interested in obtaining a modified probability distribution whose mean and standard deviation will fit the Monte Carlo results. To this end, consider a variable λ distributed accordingly with equation (22). Its mean value is equal to α_{0k} and the corresponding squared standard deviation is $2\alpha_{0k}^2/d$. It is possible to modify these quantities in order to get a better approximation to the Monte Carlo data: we replace α_{0k} and $d/2$ by two new parameters α_k and ν_k, respectively, such that

$$\langle\lambda\rangle = \alpha_k \qquad \text{and} \qquad \sigma_\lambda^2 = \frac{\alpha_k^2}{\nu_k}. \tag{30}$$

The probability distribution which verifies these conditions is a chi-squared distribution:

$$P_k(\lambda) = \frac{1}{\Gamma(\nu_k)}\frac{\nu_k}{\alpha_k}\left(\frac{\nu_k\lambda}{\alpha_k}\right)^{\nu_k-1}\exp\left(-\frac{\nu_k\lambda}{\alpha_k}\right). \tag{31}$$

In the limit $d \gg s \gg 1$ these parameters should approach: $\alpha_k \to \alpha_{0k}$ and $\nu_k \to d/2$, respectively.

An additional justification for choosing a distribution of the form (31) as the probability distribution of the inertia eigenvalues can be obtained from [6]. To fix ideas, let us consider the case of parameter α_k. In [6] a $1/d$ series expansion is obtained for the mean value of the principal inertia moments

$$\langle\lambda_k\rangle = \alpha_{0k}\left[1 + \frac{3}{4d} + \mathcal{O}\left(\frac{1}{d^2}\right)\right]. \tag{32}$$

Using (30) and (32), our new parameter α_k can be equated to the sum of a series whose zeroth-order term is the asymptotic value α_{0k}. A similar argument can be used with ν_k and $d/2$ using the series expansion for σ_λ^2 (see [6]). Therefore the distribution (22) can be regarded as a 'zeroth-order' asymptotic approximation to the more general equation (31).

The external parameters α_k and ν_k can be easily adjusted from estimators m and Δ^2 for $\langle\lambda\rangle$ and σ_λ^2 respectively—obtained from Monte Carlo data, for example. α_k and ν_k can be immediately evaluated via

$$\alpha_k = m \qquad \nu_k = \frac{m^2}{\Delta^2}. \tag{33}$$

Of course, this is not the only way to adjust the parameters; we have chosen it due to its simplicity and because the resulting distribution possess the correct mean and standard deviation, and the resulting distributions represent adequately the true ones in a large number of cases.

In figure 1 we also display the histogram corresponding to distribution (31) (large dots). It is evident that the modified distribution agrees better with the Monte Carlo data than the one of (22). In this case we obtained $\alpha_1 = 12.83$, $\nu_1 = 2.49$, using (33) with m and Δ taken from our Monte Carlo data. These figures differ significantly from the values $\alpha_{0,1} = 10.13$ and $d/2 = 1$ used in the distribution (22).

The distribution (31) seems to be adequate also for $k \neq 1$. In figure 2 we can see the case $d = 3$, $s = 100$, $k = 3$. Here the distribution (22) separates significantly from the Monte Carlo data (small dots) while our modified one gives quite good agreement for $\alpha_3 = 1.08$ and $\nu_3 = 5.95$ (large dots). In this case $\alpha_{0,3} = 1.13$, which is close to α_3: the difference between both distributions comes from the value of the parameter ν_3.

Other representative cases are shown in figure 3 where the Monte Carlo distributions (full curves) as well as the adjusted distribution (31) (dotted curves) are displayed. Here we can see that for larger values of d both histograms show an excellent agreement. Notice also how the distributions become more peaked when d grows.

It is found that when d grows the parameters α_k and ν_k tend, as they should, to their asymptotic values: for example, figure 4 shows the case $s = 100$ $k = 1$. Here $\alpha_1/\alpha_{0,1}$ (circles) and $2\nu_1/d$ (triangles) are plotted versus d. It shows up clearly that, when d grows, both curves approach the correct limit. The parameters were evaluated using (33) with estimators calculated using our Monte Carlo data. The convergence speed of these parameters to their asymptotic values, when $d \to \infty$, may depend on k.

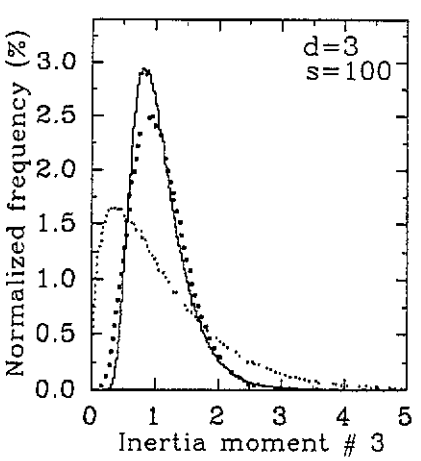

Figure 2. Same as figure 1, but for moment λ_3.

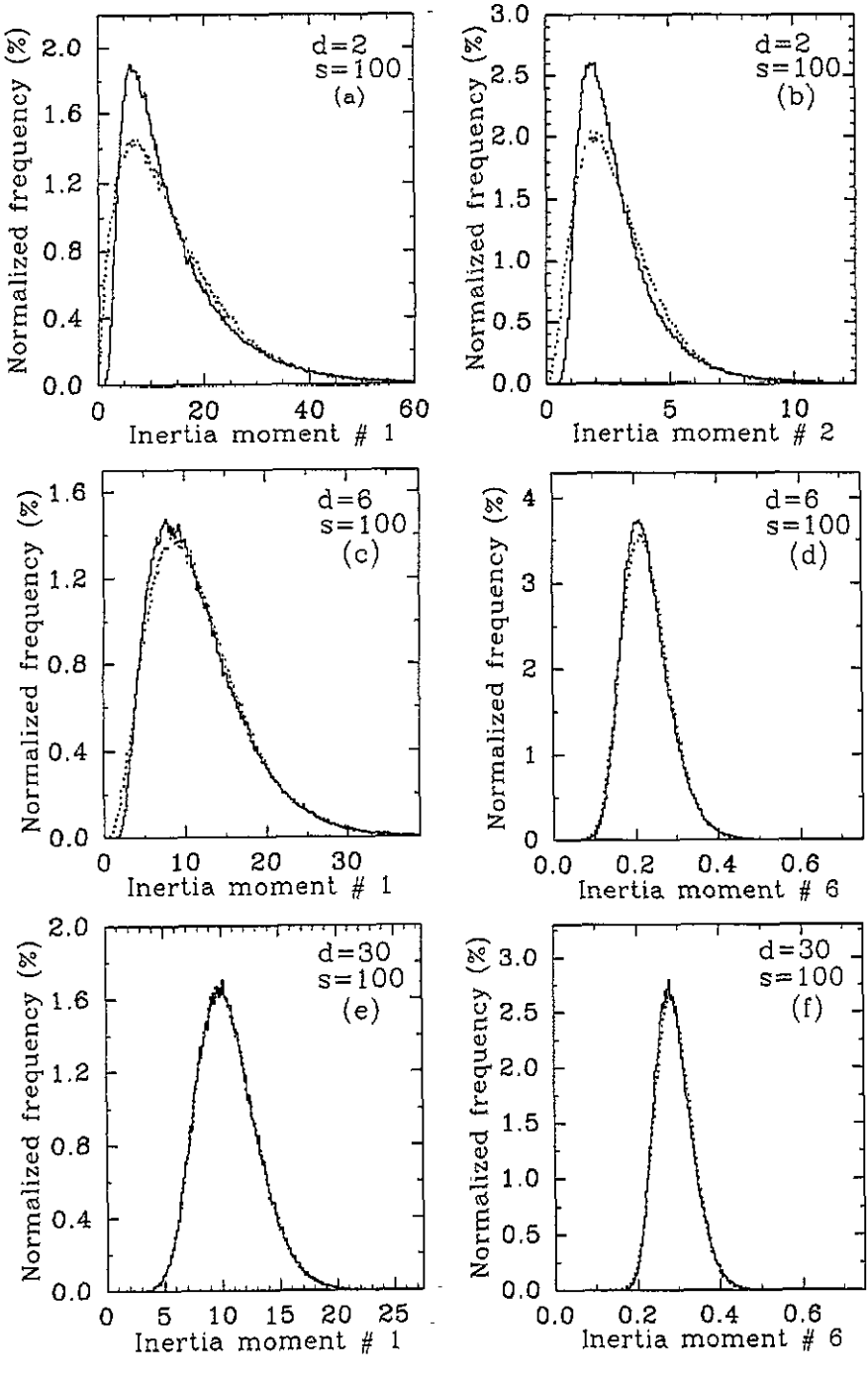

Figure 3. Frequency histograms for the inertia moments: (*a*) λ_1 and (*b*) λ_2, in the case $d = 2$, $s = 100$; (*c*) λ_1 and (*d*) λ_6, in the case $d = 6$, $s = 100$; (*e*) λ_1 and (*f*) λ_6, in the case $d = 30$, $s = 100$. The histograms in full (dotted) curves correspond to Monte Carlo data (distribution (31)). All the histograms are normalized accordingly with (28) and (29). In all cases $N' = 3 \times 10^5$; $N = 3 \times 10^5$ in cases (*a*)–(*d*) and $N = 1.5 \times 10^5$ in cases (*e*) and (*f*).

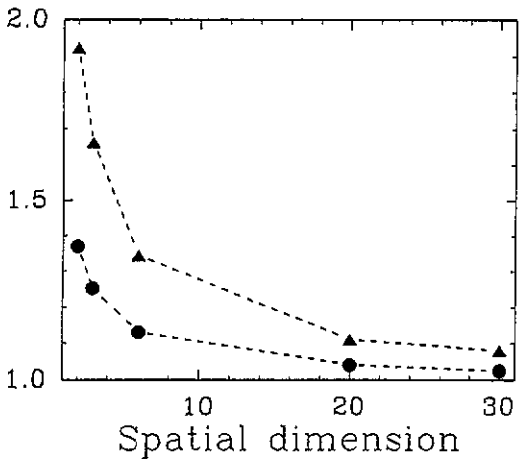

Figure 4. $\alpha_1/\alpha_{0,1}$ (circles) and $2\nu_1/d$ (triangles) plotted versus d for $s = 100$. The lines are only to guide the eye.

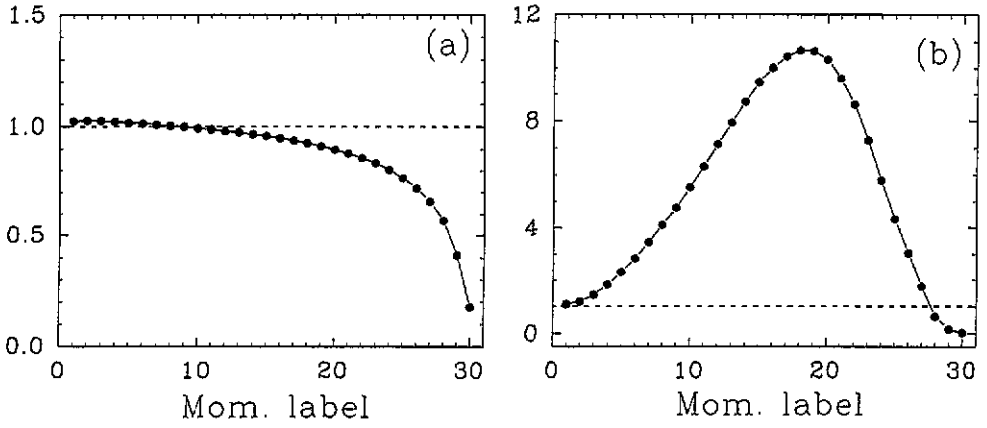

Figure 5. (*a*) α_k/α_{0k} and (*b*) $2\nu_k/d$ plotted versus the moment label k ($k = 1, \ldots, d$) for $d = 30$ and $s = 100$. The broken curves indicate the expected value of these magnitudes in the limit $d \to \infty$.

In figure 5 we display the values of (*a*) α_k/α_{0k} and (*b*) $2\nu_k/d$ versus k for $d = 30$, $s = 100$. There one can clearly see that both magnitudes are close to one just for the first two or three values of k. For larger values of k, α_k remains more or less close to its asymptotic value but ν_k separates significantly from the value $d/2$. For k near to d, both α_k and ν_k are less than α_{0k} and $d/2$, respectively, at variance with the case $k = 1$ where both magnitudes remain greater than their respective asymptotic values for all the cases that were considered. This leads to the conclusion that the distribution (22) agrees adequately with the true distribution only when the condition $d \gg s \gg k \approx 1$ holds.

We have also performed simulations for random walks with different lengths. The corresponding histograms do not differ qualitatively from the respective ones for $s = 100$. To illustrate this, we display in table 1 the mean values of λ_1 divided by s and the corresponding standard deviations. The table also includes the number N of Monte Carlo samples that were calculated in each case. The quantities in brackets indicate the error in the last two digits displayed. The errors are taken as twice the standard error of the mean,

Table 1. Mean value and standard deviation of the largest inertia moment, λ_1, divided by s, the number of steps in the random walk; tabulated for various values of d and s. The number N of Monte Carlo samples is also displayed. The quantities in brackets indicate the error in the last two digits displayed. The errors correspond to twice the standard error of the mean (equation (34)).

d	s	N ($\times 10^3$)	$\langle\lambda_1\rangle/s$	σ_{λ_1}/s
2	10^2	300	0.14028 (37)	0.1012
	10^4	300	0.13873 (37)	0.1013
	10^6	1	0.1374 (60)	0.0951
3	10^2	300	0.12829 (30)	0.08127
	10^4	200	0.12720 (37)	0.08164
6	10^2	300	0.11566 (21)	0.05759
	10^4	100	0.11453 (36)	0.05754
30	10^2	150	0.10478 (13)	0.02604
	10^4	2	0.1033 (12)	0.02619

$\Delta\mu$, related to the standard deviation via†

$$\Delta\mu = \frac{\sigma}{\sqrt{N}}. \tag{34}$$

We can see that, for fixed d, $\langle\lambda_1\rangle/s$ and σ_{λ_1}/s do not depend significantly on s. However, in all the tabulated cases s is greater (or much greater) than d. When $s \leqslant d$ the means and standard deviations may depend on s.

For s fixed and $d \to \infty$ the standard deviations go to zero accordingly with (30) (ν_k increases when d grows). Furthermore, it is not difficult to see that the standard deviation verifies approximately the law $\sigma = \text{constant}/\sqrt{d}$. This is consistent with the fact that $\nu_k \to d/2$ with $d \to \infty$.

In table 2 we present some representative values for parameters α_k and ν_k. As explained below, they were obtained using data from our Monte Carlo simulations. We can see that these parameters present a slight variation when s changes from 10^2 to 10^4. Notice that the ratio α_k/α_{0k} can also be evaluated using the series expansion (32). It is not difficult to verify that the first-order prediction $\alpha_k/\alpha_{0k} = 1 + 3/(4d)$ gives reasonable results only for very low values of k (1 or 2).

Another remarkable feature of the probability distributions of the principal inertia moments is that they possess a non-zero skewness (the mean value is not equal to the most probable one), especially for low values of d. Consider, for instance, the case $d = 3$, $s = 100$: here $\langle\lambda_1\rangle$ and the corresponding most probable value differ in about 0.7 times the standard deviation σ_{λ_1}.

This property of the distributions becomes important when trying to describe *a priori* the shape of a random walk: the most probable value is not equal to and *must* be distinguished from the mean value. Sometimes (see, for example, [7]) this difference is not clearly stated.

3.2. Other magnitudes

We have also studied the probability distribution of the asphericity and the angle Θ for several values of d and s.

† The standard error of the mean and the standard deviation of the corresponding distribution are different quantities which should not be confused (see [7] for instance). This is explained in most textbooks on statistics (see [12], for example).

Table 2. Parameter ratios α_k/α_{0k} and $2\nu_k/d$ (as obtained from (33) with data coming from our Monte Carlo simulations; tabulated for representative values of d, s and k. In all cases the number N of Monte Carlo samples is greater than 10^5.

d	s	k	α_k/α_{0k}	$2\nu_k/d$
2	10^2	1	1.3707	1.9227
	10^2	2	1.1083	3.2194
	10^4	1	1.3690	1.8769
	10^4	2	1.0972	3.1012
3	10^2	1	1.2535	1.6614
	10^2	2	1.1361	2.7559
	10^2	3	0.9480	3.9698
	10^4	1	1.2553	1.6184
	10^4	2	1.1323	2.6596
	10^4	3	0.9396	3.7764
6	10^2	1	1.1301	1.3443
	10^2	2	1.1060	2.0109
	10^2	6	0.7951	5.3292
30	10^2	1	1.0238	1.0791
	10^2	6	1.0135	2.8277
	10^2	30	0.1765	0.0319

In figure 6 we display the probability distribution for the asphericity in several representative cases. The full curves correspond to the Monte Carlo data, while the dotted ones correspond to frequency histograms built using an auxiliary Monte Carlo process to simulate the probability distribution for the asphericity (equation (26)), taking the combined probability (27) as a product of distributions of the form (31). This auxiliary process consists in the generation of a large number of sets of random numbers $z_1, \ldots, z_r$ ($r = \min(d, s)$), each one distributed accordingly with the distribution (31), with the parameters corresponding to $\lambda_1, \ldots, \lambda_r$, respectively. Then each set of numbers is used as input for (13) (taking $\lambda_{r+1} = 0, \ldots, \lambda_d = 0$ when necessary) in order to obtain a sample for the asphericity. The set of samples is processed to build a histogram, which constitutes a numerical estimation of the distribution (26). All histograms in figure 6 were normalized as explained in section 3.1.

We can observe broad distributions, especially for low values of d, with the following remarkable characteristic: for $d = 2$ the probability distributions for the asphericity are different from zero at the origin, whereas they vanish at this point for all $d \geqslant 3$. In other words, there is a finite probability of generating a random walk with A very close to zero just for $d = 2$ while for $d \geqslant 3$ this probability vanishes. Similar contrasts between the $d = 2$ and the $d > 2$ cases can be found when considering other properties of random walks, like the probability of returning to the origin [13], for example.

When d grows, the distributions become more peaked and approach the limit of infinite d already described in section 2.1. For example, figure 7 shows the case $d = 30$, $s = 10$. We can clearly see the peak at the most probable value, which is approximately equal to the limit of A for $d \to \infty$ (equation (19)). Notice that the distributions possess a non-zero skewness, especially for low values of d (see the related remark in the previous section).

For increasing s the distributions do not alter significantly, as can be inferred from the data of table 3. In this table the mean values of A and Θ were tabulated for several values of d and s. The number N of Monte Carlo samples used is also included. The quantities in

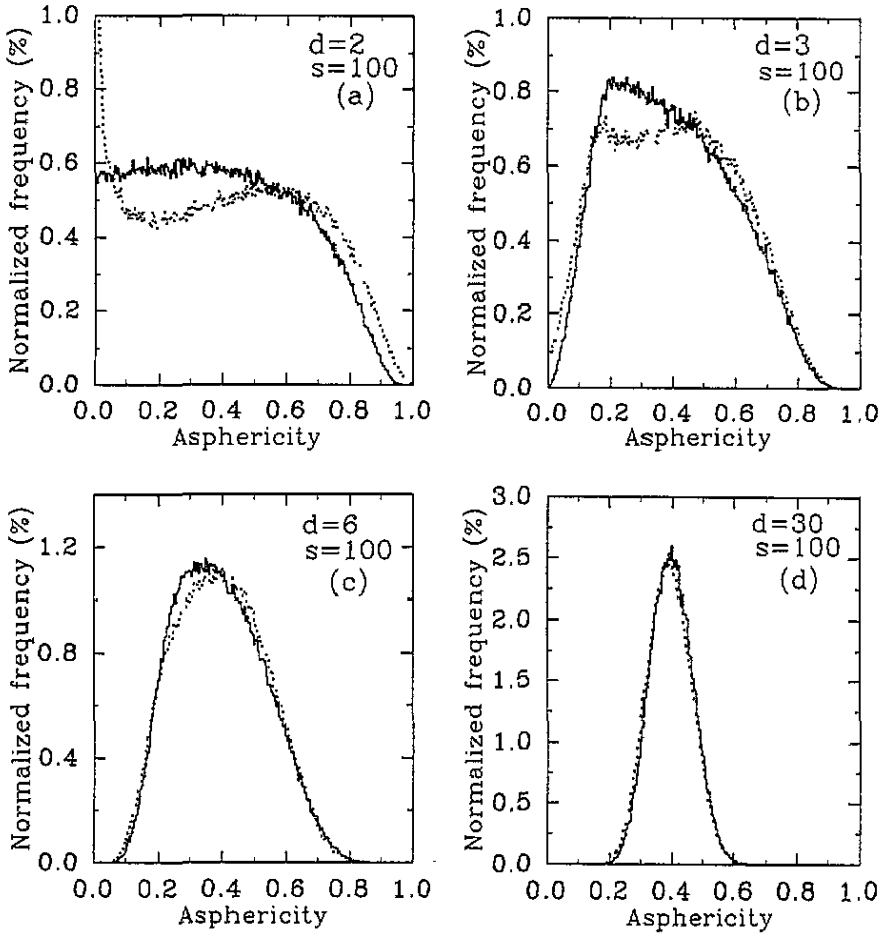

Figure 6. Frequency histograms for the asphericity A, in the cases $s = 100$ and d equal to (a) 2, (b) 3, (c) 6 and (d) 30. The full curves correspond to the results of the Monte Carlo simulations while the dotted ones correspond to the auxiliary Monte Carlo simulation of distribution (26) for the asphericity. (The combined probability distribution Q is taken as products of distributions of the form of equation (31).) The dotted histogram in the $d = 2$ case represents the distribution (38). All the histograms are normalized accordingly with (28) and (29). In all cases $N' = 3 \times 10^5$; $N = 3 \times 10^5$ in cases (a)–(c) and $N = 10^5$ in case (d). The point $A = 0.4$ corresponds to the asymptotic value of (19).

brackets indicate the error in the last two digits displayed (see equation (34). For d fixed, both the mean value $\langle A \rangle$ and standard deviation σ_A tend to a definite value for $s \to \infty$. For s fixed and increasing d the standard deviations diminish, as expected. It is not difficult to verify that the standard deviations vary with the law

$$\sigma \cong \text{constant} \times d^{-w} \tag{35}$$

with $w \approx 0.45$.

The Monte Carlo histogram of figure 6(b) should be compared with a similar one presented in [8] for Gaussian random walks of length $s = 1024$ [8, figure 1]. It is possible to see that there is a good agreement between both distributions. A qualitatively similar

Table 3. Mean value and standard deviation of the asphericity, A, and the angle between the principal axis of inertia and the end-to-end vector, Θ, together with the corresponding standard deviations, tabulated for various values of d and s. The number N of Monte Carlo samples is also displayed. The quantities in brackets indicate the error in the last two digits displayed. The errors correspond to twice the standard error of the mean (equation (34)).

d	s	N ($\times 10^3$)	$\langle A \rangle$	σ_A	$\langle \Theta \rangle$	σ_Θ
2	10	300	0.385 63(86)	0.235 1	0.413 1(16)	0.440 0
	10^2	300	0.393 81(86)	0.235 4	0.411 2(14)	0.387 3
	10^3	300	0.395 35(86)	0.236 3	0.412 2(14)	0.381 0
	10^4	300	0.395 86(86)	0.236 0	0.413 0(14)	0.380 5
	10^5	300	0.396 16(86)	0.236 0	0.412 1(14)	0.380 7
3	10	300	0.390 07(69)	0.189 0	0.425 1(13)	0.361 4
	10^2	300	0.392 36(69)	0.189 2	0.469 0(13)	0.350 3
	10^3	300	0.394 21(69)	0.189 7	0.475 9(13)	0.352 5
6	10	300	0.399 00(51)	0.138 3	0.408 56(94)	0.256 6
	10^2	300	0.392 50(51)	0.139 9	0.480 61(97)	0.266 1
	10^3	300	0.393 87(51)	0.140 0	0.487 79(98)	0.267 6
30	*10*	*100*	*0.410 94(41)*	*0.065 35*	*0.362 97(55)*	*0.086 75*
	10^2	100	0.397 73(43)	0.067 96	0.445 98(58)	0.091 62
	10^3	20	0.398 14(96)	0.067 89	0.454 6(13)	0.092 86

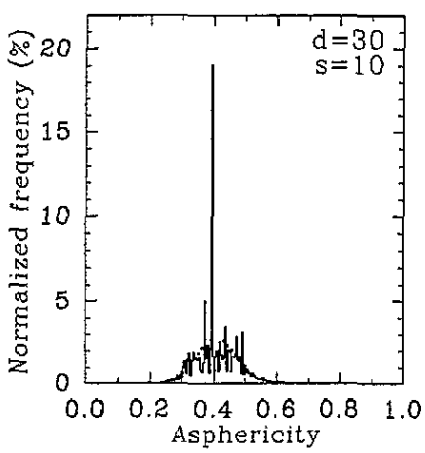

Figure 7. Same as figure 6, but for the case $d = 30$, $s = 10$, $N' = 3 \times 10^5$, $N = 10^5$.

histogram is obtained from the simulation of self-avoiding Gaussian random walks [8, figure 8]. These two facts reinforce the idea of universality in the shape properties of random walks [8].

When looking at the dotted histograms which correspond to the distribution (26) in the case of the asphericity, we can see that in general there is an acceptable, but not complete agreement. This shows up clearly in the case $d = 2$ where both distributions differ significantly and in the case $d = 30$, $s = 10$, where the 'theoretical' distribution does not predict the peak at $A \cong 0.4$.

We have analysed the reasons for such behaviour to conclude that there are two main sources for the disagreements: (i) the independence assumption of (27) does not hold exactly, especially for low d. Observe, for example, that in the $d = 2$ case mentioned, the distributions for the inertia moments (figures 3(a) and (b)) present a good agreement with

the true ones, while the combined probability (equation (27)) gives a manifestly discordant distribution for A (figure 6(a)). (ii) When $s \approx d$ and especially when $s \leqslant d$, the probability distributions for λ_k with $k \cong r$ are discrete (just a few isolated values of λ_k possess a non-zero probability) and it remains clear therefore that this kind of distributions cannot be adequately described in terms of a chi-squared distribution.

The fact that the combined probability distribution of (27) does not hold exactly can be understood simply. When defining the principal inertia moments as the eigenvalues of the inertia matrix (3), the assumption $\lambda_1 \geqslant \lambda_2 \geqslant \cdots \geqslant \lambda_d$ was made. This imposes, by construction, the constraint

$$Q(\lambda_1, \cdots, \lambda_r) \begin{cases} \geqslant 0 & \text{if} \quad \lambda_1 \geqslant \lambda_2 \geqslant \cdots \geqslant \lambda_r \\ = 0 & \text{otherwise}. \end{cases} \tag{36}$$

Apart from this, there are other restrictions that can be imposed *a priori*. For example, from (4) and (11) it follows that Q must be zero when $\lambda_1 + \ldots + \lambda_r > S^2_{\max}$. These conditions, which imply that the different inertia moments are not independent, are certainly not verified by the distribution (27), especially in the cases when the different distributions $P_k(\lambda)$ overlap significantly among themselves.

Looking at the distributions of figures 1–3 one can notice that the overlaps between distributions are significant for low values of d, and tend to zero as d increases (see the different scales used in the horizontal axes). Therefore, one can clearly understand the reason for the failure of the distribution for the asphericity described previously (figure 6(a)).

In this particular case of two dimensions, it is possible to evaluate analytically $P_A(x)$. The details of the calculations can be found in the appendix, and the result is

$$P_A(x) = \frac{1}{\sqrt{x}} \int_0^\infty \left[\frac{Q(\lambda, \lambda\beta^{-1}(x))}{(1-\sqrt{x})^2} + \frac{Q(\lambda, \lambda\beta(x))}{(1+\sqrt{x})^2} \right] \lambda \, \mathrm{d}\lambda \tag{37}$$

where $x \in [0, 1]$ and $\beta(x)$ is defined in (A6). Notice that $\beta(x) \leqslant 1$ and then if the constraint (36) is in effect, the first term in the integrand vanishes identically, and the distribution reduces to (A8).

If one uses the distribution (27) in (37), then $P_A(x)$ takes the form (see appendix)

$$P_A(x) = \frac{\Gamma(\nu_1+\nu_2)}{\Gamma(\nu_1)\Gamma(\nu_2)} \left(\frac{\nu_1}{\alpha_1}\right)^{\nu_1} \left(\frac{\nu_2}{\alpha_2}\right)^{\nu_2} \frac{1}{\sqrt{x}(1-\sqrt{x})^2} \times \left\{ \frac{[\beta(x)]^{1-\nu_2}}{\left[\frac{\nu_1}{\alpha_1} + \frac{\nu_2}{\alpha_2\beta(x)}\right]^{\nu_1+\nu_2}} + \frac{[\beta(x)]^{1+\nu_2}}{\left[\frac{\nu_1}{\alpha_1} + \frac{\nu_2\beta(x)}{\alpha_2}\right]^{\nu_1+\nu_2}} \right\}. \tag{38}$$

It is not difficult to see that, for $x \to 0$, $P_A(x)$ can be written as follows:

$$P_A(x) = \left\{ \frac{2\Gamma(\nu_1+\nu_2)}{\Gamma(\nu_1)\Gamma(\nu_2)} \frac{\left(\frac{\nu_1}{\alpha_1}\right)^{\nu_1}\left(\frac{\nu_2}{\alpha_2}\right)^{\nu_2}}{\left(\frac{\nu_1}{\alpha_1}+\frac{\nu_2}{\alpha_2}\right)^{\nu_1+\nu_2}} \right\} \frac{1}{\sqrt{x}} + \mathcal{O}(x^0). \tag{39}$$

This equation tells us that $P_A(x)$ is divergent when $x \to 0$. In figure 6(a) we can clearly see this divergence in the dotted histogram (which was obtained after an auxiliary Monte Carlo simulation of distribution (38)). The comparison of the distribution (38) with the auxiliary Monte Carlo simulation provided us with a powerful check for the procedure.

We can describe in similar terms the behaviour of the probability distribution for the angle Θ. In figure 8 we can see some examples of such distributions. We display only the Monte Carlo data, since a probability distribution for Θ cannot be obtained from the probability distributions (22) or (31).

Notice that for low values of d all the distributions are broad, present a non-zero skewness and vanish (do not vanish) for $\Theta = 0$ when $d \geqslant 3$ ($d = 2$).

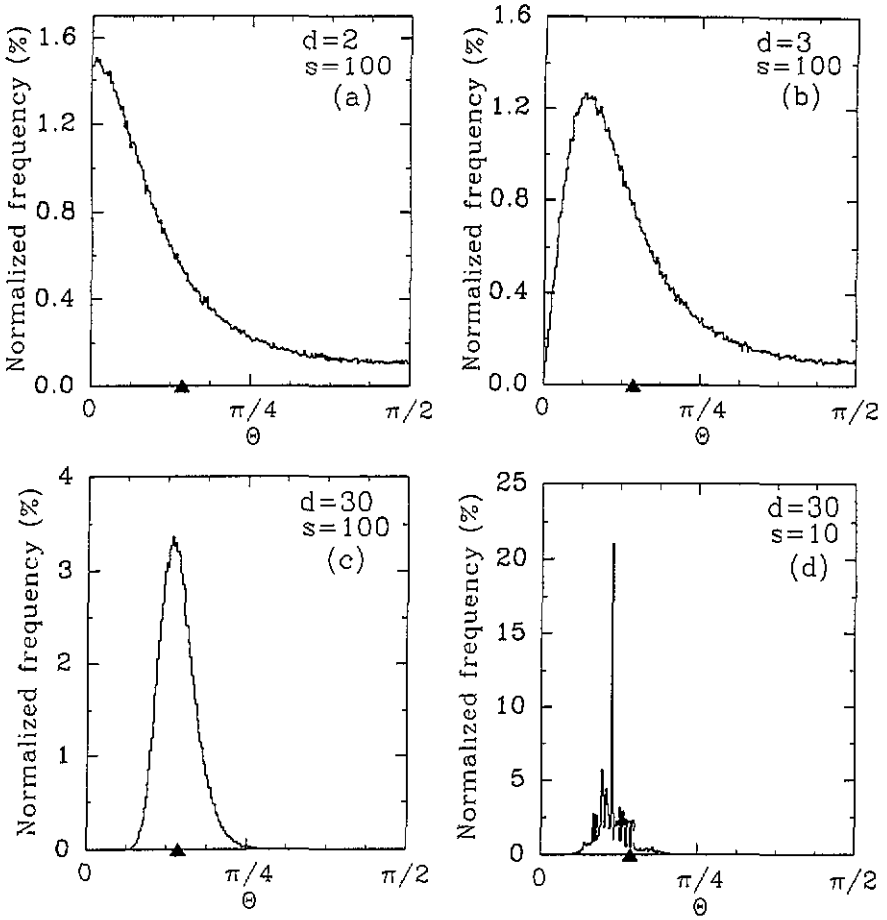

Figure 8. Frequency histograms for the first quadrant angle Θ between the principal inertia axis (associated with moment λ_1) and the end-to-end vector r_s, in the cases: (*a*) $d = 2$, $s = 100$; (*b*) $d = 3$, $s = 100$; (*c*) $d = 30$, $s = 100$; (*d*) $d = 30$, $s = 10$. All the histograms are normalized accordingly with equations (28) and (29). In cases (*a*), (*b*) $N = 3 \times 10^5$, and in cases (*c*), (*d*) $N = 10^5$. The triangle in the Θ axis indicates the asymptotic value of (21).

For large d the probability distribution peaks around a central value. For example, in figure 8(*d*) we present the case $d = 30$, $s = 10$. The peak is very near to the point which corresponds to the limit $d \to \infty$ (equation (20)). This effect can also be detected analysing the corresponding data of table 3: the standard deviations verify a law of the form (35) with $w \approx 0.55$.

In connection with the study of angular correlations of [7], the results already described lead to the conclusion that in the case of low dimension there are no strong correlations between the end-to-end vector and the principal inertia axis, independently of the length s of the random walks. For instance, in the $d = 3$ case, the relative fluctuation $\sigma_\Theta/\langle\Theta\rangle$ is approximately equal to 75% for all the lengths considered (see table 3). Similar conclusions apply to the case of other angles [7] used to study correlations.

4. Conclusions and final remarks

We have performed a detailed study of the behaviour of the probability distribution of several magnitudes defined on open random walks in discrete spaces with coordination number equal to four, with the main purpose of analysing the shape of these objects. The selected magnitudes, largely used in previous related works [3–10], are: the principal inertia moments, $\lambda_1, \ldots, \lambda_d$, the asphericity, A, and the angle Θ between the principal inertia axis and the end-to-end vector.

In general, the distributions spread widely around a single most probable value. In many cases this value differs significantly from the corresponding mean value due to a non-zero skewness of the respective distributions. The distributions become more peaked and symmetric when the space dimension d grows, especially for small values of the length of the random walk, in agreement with the analytic results that can be obtained in the limit $d \gg s$, discussed in section 2.

In the case of random walks of large length ($s \gg d$), the shape of the distributions does not depend significantly on s. This implies that the relative fluctuations (ratio between standard deviation and mean value) of the corresponding magnitudes do not vary—and therefore do not go to zero—in the limit of very large lengths. This fact leads to the conclusion that the mean values of any related quantity cannot be considered complete descriptors of the shapes of a given set of random walks. For the mean values to be useful by themselves, the existence of a 'thermodynamical' limit for the system is necessary (such a limit exists when fluctuations go to zero for very large sizes of the system).

This conclusion applies not only to the variables studied here but also to other magnitudes commonly introduced to describe the shape of random walks, i.e. inertia moment ratios, etc [3, 7].

The analysis of the probability distributions for the principal inertia moments shows that they can be fitted to chi-squared distributions (equation (31)). These distributions approach asymptotically the distribution of (22) [6] in the limit $d \gg s \gg 1$, and when the moment label k of λ_k is not large ($k \ll r$, $r = \min(d, s)$). When these conditions are not verified, distributions (31) and (22) may differ significantly (see section 3.1).

We also studied the probability distributions for A and Θ, finding that both magnitudes present, in most of the considered cases, wide distributions with non-zero skewness, especially for low values of d. For large d the distributions become more peaked around a certain value, which corresponds to the limit for $d \to \infty$ discussed in section 2. For fixed s, the corresponding standard deviations diminish in the form of (35) when d grows.

We used the probability distribution (31) to build the combined probability distribution for the entire set of non-vanishing inertia moments $\lambda_1, \ldots, \lambda_r$, and by means of an auxiliary Monte Carlo process we evaluated the probability distribution for the asphericity and compared it with the results of our simulations. There is an acceptable concordance between both distributions in all the cases considered, with the exceptions of the cases with very low d ($d = 2, 3$), and $d = 30$, $s = 10$. One of the reasons for such disagreements is that the independence hypothesis assumed when constructing the combined probability distribution for the entire set, $\lambda_1, \ldots, \lambda_r$, of non-vanishing inertia moments (equation (27)), is not verified exactly, especially for low values of d (see section 3.2).

Finally, it is worthwhile mentioning that it will be interesting to perform the analytic counterpart of this mainly numerical analysis. There are many results in our paper which prompt analytic calculations. For example, evaluation of parameters α_k and ν_k in distribution (31), corrections to this distribution to fit the true one, analysis of combined probability distributions, etc. We leave this for a future publication.

Acknowledgments

We are indebted to L N Epele, C A García Canal and H Fanchiotti for useful discussions and suggestions. We are also indebted to J L Alessandrini for letting us know about many of the references and for useful hints. Finally, we are grateful to Fundación Antorchas of Argentina for its help. This work was partially supported by the Consejo Nacional de Investigaciones Científicas y Técnicas of Argentina.

Appendix. The asphericity in the $d = 2$ case

The asphericity (12) of two-dimensional objects is given by

$$A = \left(\frac{\lambda_1 - \lambda_2}{\lambda_1 + \lambda_2}\right)^2 . \tag{A1}$$

Then, if $Q(\lambda_1, \lambda_2)$ is the combined probability distribution of the eigenvalues λ_1 and λ_2, the probability distribution of A is given by (see equation (26))

$$P_A(x) = \int_0^\infty d\lambda_1 \int_0^\infty d\lambda_2 \, \delta(x - A)\, Q(\lambda_1, \lambda_2) \qquad 0 \leqslant x \leqslant 1 . \tag{A2}$$

The integral in λ_2 can be evaluated after taking into account that

$$\delta(x - f(y)) = \sum_{\substack{\text{all } y_k \\ \text{such that } x = f(y_k)}} \frac{1}{|f'(y_k)|}\delta(y - y_k) . \tag{A3}$$

In our case, the equation $x - A(\lambda_1, \lambda_2) = 0$ possesses two solutions, namely

$$\lambda_2 = \frac{1 - \theta\sqrt{x}}{1 + \theta\sqrt{x}}\lambda_1 \qquad \theta = \pm 1 . \tag{A4}$$

The derivative $\partial A/\partial\lambda_2$ at these points is given by

$$\left.\frac{\partial A}{\partial \lambda_2}\right|_{x=A} = -\frac{\theta\sqrt{x}(1 + \theta\sqrt{x})^2}{\lambda_1} . \tag{A5}$$

Then, using (A3) and (A5), the properties of the δ distribution, and defining

$$\beta(x) = \frac{1 - \sqrt{x}}{1 + \sqrt{x}} \tag{A6}$$

one can write (A2) as follows:

$$P_A(x) = \frac{1}{\sqrt{x}} \int_0^\infty \left[\frac{Q(\lambda, \lambda\beta^{-1}(x))}{(1 - \sqrt{x})^2} + \frac{Q(\lambda, \lambda\beta(x))}{(1 + \sqrt{x})^2}\right] \lambda \, d\lambda . \tag{A7}$$

When the constraint (36) is in effect, and as $0 \leqslant \beta(x) \leqslant 1$ for all x in $[0, 1]$, the first term in the integrand is zero. Consequently, in this case P_A reduces to

$$P_A(x) = \frac{1}{\sqrt{x}(1 + \sqrt{x})^2} \int_0^\infty Q(\lambda, \lambda\beta(x))\, \lambda \, d\lambda . \tag{A8}$$

On the other hand, if the combined probability distribution Q is of the form (27), $Q(\lambda_1, \lambda_2) = P_1(\lambda_1)P_2(\lambda_2)$, equation (A7) must be used to evaluate $P_A(x)$. If P_1 and P_2 are

given by (31), the resulting distribution can be evaluated straightforwardly. In fact, from (31),

$$Q(\lambda, \lambda\beta^{\pm 1}(x)) = \frac{1}{\Gamma(\nu_1)\Gamma(\nu_2)} \left(\frac{\nu_1}{\alpha_1}\right)^{\nu_1} \left(\frac{\nu_2}{\alpha_2}\right)^{\nu_2} [\beta(x)]^{\pm(\nu_2-1)} \lambda^{\nu_1+\nu_2-2} \times \exp\left[-\left(\frac{\nu_1}{\alpha_1} + \frac{\nu_2}{\alpha_2}\beta^{\pm 1}(x)\right)\lambda\right]. \tag{A9}$$

Using these distributions in (A7) we obtain

$$P_A(x) = \frac{1}{\Gamma(\nu_1)\Gamma(\nu_2)} \left(\frac{\nu_1}{\alpha_1}\right)^{\nu_1} \left(\frac{\nu_2}{\alpha_2}\right)^{\nu_2} \frac{1}{\sqrt{x}(1-\sqrt{x})^2} \times \int_0^\infty \left[\beta^{1-\nu_2} e^{-(\nu_1/\alpha_1+\nu_2/\alpha_2\beta(x))\lambda} + \beta^{1+\nu_2} e^{-(\nu_1/\alpha_1+\nu_2/\alpha_2\beta(x))\lambda}\right] \lambda^{\nu_1+\nu_2-1} \, d\lambda . \tag{A10}$$

The integral can be easily evaluated considering that

$$\int_0^\infty t^{u-1} e^{-vt} \, dt = \frac{\Gamma(u)}{v^u} \qquad v \geqslant 0. \tag{A11}$$

Then,

$$P_A(x) = \frac{\Gamma(\nu_1+\nu_2)}{\Gamma(\nu_1)\Gamma(\nu_2)} \left(\frac{\nu_1}{\alpha_1}\right)^{\nu_1} \left(\frac{\nu_2}{\alpha_2}\right)^{\nu_2} \frac{1}{\sqrt{x}(1-\sqrt{x})^2} \times \left\{ \frac{[\beta(x)]^{1-\nu_2}}{\left[\frac{\nu_1}{\alpha_1} + \frac{\nu_2}{\alpha_2\beta(x)}\right]^{\nu_1+\nu_2}} + \frac{[\beta(x)]^{1+\nu_2}}{\left[\frac{\nu_1}{\alpha_1} + \frac{\nu_2\beta(x)}{\alpha_2}\right]^{\nu_1+\nu_2}} \right\}. \tag{A12}$$

As explained in section 3.2, this distribution separates significantly from the true one, especially for small values of x.

References

[1] de Gennes P G 1979 *Scaling Concepts in Polymer Physics* (Ithaca, NY: Cornell University Press)
[2] Fernandez R, Frölich J and Sokal A D 1992 *Random Walks, Critical Phenomena and Triviality in Quantum Field Theory* (Berlin: Springer)
[3] Šolc K and Stockmayer W H 1971 *J. Chem. Phys.* **54** 2756
Šolc K 1971 *J. Chem. Phys.* **55** 335
[4] Šolc K and Gobush W 1974 *Macromolecules* **7** 814
[5] Rudnick J, Beldjenna A and Gaspari G 1987 *J. Phys. A: Math. Gen.* **20** 971
[6] Gaspari G, Rudnick J and Beldjenna A 1987 *J. Phys. A: Math. Gen.* **20** 3393
[7] Bruns W 1992 *Makromol. Chem. Theory. Simul.* **1** 287
[8] Jagodzinski O, Eisenriegler E and Kremer K 1992 *J. Physique I* **2** 2243
[9] Gaspari G, Rudnick J and Beldjenna A 1993 *J. Phys. A: Math. Gen.* **26** 1
[10] Rudnick J and Gaspari G 1986 *J. Phys. A: Math. Gen.* **19** L191
[11] Eichinger B E 1980 *Macromolecules* **13** 1
[12] Lohness P R and Cooley W W 1968 *Introduction to Statistical Procedures* (New York: Wiley)
[13] Bender C M, Boettcher S and Mead L R 1994 *J. Math. Phys.* **35** 368