The shape of self-avoiding walks

1996 J. Phys. A: Math. Gen. 29 5455

(http://iopscience.iop.org/0305-4470/29/17/019)

J. Phys. A: Math. Gen. **29** (1996) 5455–5473. Printed in the UK

The shape of self-avoiding walks

S J Sciutto†

Laboratorio de Física Teórica, Departamento de Física, Universidad Nacional de La Plata, CC 67-1900 La Plata, Argentina

Received 13 February 1996, in final form 14 May 1996

Abstract. An exhaustive analysis of the shape properties of discrete self-avoiding random walks is presented. The dependence of the main parameters of the probability distributions from the length of the walk and the differences between these distributions for conventional and self-avoiding walks are studied. By means of high-precision Monte Carlo simulations it is shown that the characteristics of the shape of self-avoiding random walks, when regarded as a function of the walk length, present the expected asymptotic behaviour. The differences with conventional random walks depend upon the observable considered: the probability distributions of the principal inertia eigenvalues of self-avoiding walks spread around the most probable value more widely than the corresponding distributions for unrestricted walks, while in the cases of the asphericity or ratios of inertia eigenvalues, the distributions for self-avoiding walks are somewhat more peaked than their counterparts for conventional random walks. Analytical expressions for the probability distributions of inertia moment ratios and the two-dimensional asphericity are given. For common random walks there is an excellent agreement between these analytical distributions and the Monte Carlo data. In this work it is established that this concordance is maintained if the same analytical distributions are applied to the self-avoiding case. This means that within the precision of the simulations the functional form of the mentioned distributions does not vary when passing from conventional to self-avoiding walks.

1. Introduction

Both the conventional random walk (RW) and the self-avoiding walk (SAW) represent well known objects appearing in diverse areas of physics such as quantum field theory [1] and the theory of polymer molecules [2]. One interesting feature of both SAWs and RWs is that their average shape is not spherical [3–10]. This property is of interest when analysing some polymeric fluids [8, 11], and/or the effect of the concentration [12] or interactions [13] on the mean dimensions of single polymer chain models.

In previous papers [3, 4] we presented a comprehensive study of the shapes of discrete RWs in spaces of arbitrary dimension, and it is the main purpose of the present work to report the results of a similar exhaustive analysis in the case of discrete SAWs.

There are previous studies about shape properties of SAWs: in [5] the mean values of the principal inertia moments of both RWs and SAWs are analysed; in [8] the asphericity of two- and three-dimensional Gaussian SAWs is studied, together with other properties of open and ring walks; and in [9] the existence of correlations between the principal axes of inertia and the end-to-end vector is investigated. Other works treating the problem of the shape properties of walks (but not necessarily referring to SAWs) are briefly reviewed in [3].

† E-mail address: sciutto@venus.fisica.unlp.edu.ar

0305-4470/96/175455+19$19.50

With the exception of [3, 8], all the works cited in the previous paragraph deal primarily with mean values of observables; in general there is no explicit mention either of other statistical parameters like standard deviations, etc, or of probability distributions of observables.

Our study is mainly devoted to the analysis of probability distributions of several shape-describing quantities. In most cases these distributions are obtained from high-precision Monte Carlo simulations. The SAWs are generated using the pivot algorithm [14, 15] which, at present, is considered to be one of the most efficient procedures for generating lattice SAWs.

The scaling properties of some quantities, as long as the size of the walks goes to infinity, is analysed using the data coming from our simulations. The SAW scaling properties have been extensively studied [1, 16–21] and thoroughly reviewed in [22]. However, all these works refer to global size properties like the well known radius of gyration, for example, and no explicit reference is made to shape-describing magnitudes like the ones studied here. Therefore, we considered it of interest to perform direct tests of the scaling behaviour of these shape quantities, and decided to present the results in this paper.

Such analysis put into evidence that the main properties of the shape distributions, such as means or standard deviations, present only small variations when the size of the SAWs is changed, and that these variations are due to the fact that the higher order terms in the asymptotic series for the corresponding observables are not completely negligible. The sign and magnitude of these 'corrections to scaling' are concordant with the results presented in [22].

Given that the shape properties do not substantially vary when the size of the walks goes to infinity, it is sufficient to study them at a convenient fixed size in order to obtain an adequate knowledge of their behaviour.

This fact is taken into account in our work when analysing the probability distributions of many quantities, namely, the principal inertia moments, some ratios of them, the so-called asphericity [3, 4] and the first quadrant angle between the principal inertia axis and the end-to-end vector [3, 9]. In all cases, the probability distributions of both SAWs and RWs are obtained and compared, in order to put into evidence the differences existing between these two kinds of objects.

In the case of the ratios of inertia moments, the validity of the analytic probability distribution introduced in [4] for RWs, is successfully checked for the SAW case. The derived analytical probability distribution for the two-dimensional asphericity [4] is also checked. In both cases there is an excellent agreement between the analytical and true distributions. The results presented in [4] and in the present work indicate that within the precision of the simulation, the functional form of the mentioned probability distributions is the same for both RWs and SAWs.

This paper is organized as follows. Section 2 contains the definitions of most of the observables needed to adequately describe the shape of RWs and/or SAWs. Section 3 is devoted to the description of the numerical algorithm used to generate the SAWs. Section 4 is dedicated to a detailed exposition of the results of the Monte Carlo simulations, including comparisons of our data with previous works whenever possible, and analysing the analytical probability distributions of the inertia moment ratios and the two-dimensional asphericity. Finally, in section 5 we give some concluding remarks.

2. Definitions

Both conventional or self-avoiding random walks of s steps in a d-dimensional space can be defined as sets of $s + 1$ d-dimensional vectors $\boldsymbol{r}_\alpha$, $\alpha = 0, \ldots, s$, which represent the

positions within each single walk [3]. With no loss of generality we will take $\boldsymbol{r}_0 = \mathbf{0}$. We also introduce the step vectors

$$\boldsymbol{r}_\alpha = \boldsymbol{r}_{\alpha-1} + \boldsymbol{\varepsilon}_\alpha \qquad \alpha = 1, \ldots, s. \tag{1}$$

For discrete unrestricted walks with coordination number $2d$, each one of the vectors $\boldsymbol{\varepsilon}_\alpha$ may be any one of the unitary vectors $\pm\boldsymbol{e}_1, \pm\boldsymbol{e}_2, \ldots, \pm\boldsymbol{e}_d$, where $B = \{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_d\}$ is an orthonormal basis of the d-dimensional space.

For SAWs we must impose the additional self-avoidance constraint

$$\boldsymbol{r}_\alpha \neq \boldsymbol{r}_\beta \qquad \text{for all } \alpha \neq \beta. \tag{2}$$

To obtain a convenient measure of the 'size' and/or the shape of a walk, we define the following quantities [3]:

(i) The *centre of mass*:

$$\boldsymbol{r}_{\mathrm{CM}} = \frac{1}{s+1}\sum_{\alpha=1}^{s} \boldsymbol{r}_\alpha. \tag{3}$$

(ii) The *inertia matrix*:

$$T_{ij} = \frac{1}{s+1}\sum_{\alpha=0}^{s}(x_{i\alpha} - x_{\mathrm{CM}i})(x_{j\alpha} - x_{\mathrm{CM}j}) \qquad 1 \leqslant i \leqslant d, 1 \leqslant j \leqslant d. \tag{4}$$

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, $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_d$.

These quantities are used to define other derived magnitudes which are adequate to describe the geometric properties of the walk. Within this paper we are going to use the following ones.

(i) The *radius of gyration*, defined as the trace of the matrix T [3]:

$$S^2 = \mathrm{Tr}\, T = \sum_{k=1}^{d} \lambda_k. \tag{5}$$

The radius of gyration can be taken as a measure of the 'spherical volume' occupied by the walk.

(ii) Ratios of inertia moments:

$$R_{ij} = \frac{\lambda_i}{\lambda_j} \qquad i \neq j,\ \lambda_j \neq 0. \tag{6}$$

For convenience and with no loss of generality, we impose $i < j$ [4]. With this restriction we have $\lambda_i \leqslant \lambda_j$ and therefore $0 \leqslant R_{ij} \leqslant 1$.

(iii) The *asphericity*:

$$A = \frac{1}{d-1}\left[d\left(\sum_{i=1}^{d}\lambda_i^2\right)\left(\sum_{i=1}^{d}\lambda_i\right)^{-2} - 1\right]. \tag{7}$$

This quantity takes values between 0 and 1. $A = 0$ (1) corresponds to a perfectly spherical (rod) shape [3, 4].

(iv) The first quadrant angle between the path ends and the inertia axis corresponding to the largest eigenvalue λ_1 [3, 9]:

$$\cos\Theta = \frac{\boldsymbol{r}_s \cdot \boldsymbol{u}_1}{\|\boldsymbol{r}_s\|\|\boldsymbol{u}_1\|}. \tag{8}$$

One of the common ways of studying the geometric or shape properties of both conventional or self-avoiding walks is by means of a computer simulation, where statistically

independent walks of a given length s are generated. In general, this provides a set of walks with different shapes, which can be used to build frequency histograms for quantities such as the asphericity and the principal inertia moments. The same data can be used to evaluate other quantities such as mean values, standard deviations, etc. The frequency histogram data, adequately normalized, can be used to numerically approximate the probability distribution of the corresponding observable. The details of such calculations are given in the following sections.

3. The Monte Carlo algorithm

To perform, by means of a Monte Carlo simulation, a statistical analysis about any given property of a random walk, one must first provide a set of N statistically independent samples generated at random. This requirement can easily be fulfilled in the case of conventional RWs [3, 4]. On the other hand, to randomly generate a set of statistically independent SAWs is a task that requires some special considerations and the use of more involved algorithms.

First, it is necessary to define more rigorously the concept of 'a SAW generated at random' introduced in the preceding paragraph. To this end, let us define the set Ω (Ω_0) of all conventional (self-avoiding) d-dimensional RWs of length s. Clearly $\Omega_0 \subset \Omega$ since any SAW is also a common RW. Let K (K_0) be the number of elements of Ω (Ω_0); we have $K_0 < K$. We can then define the 'randomness' or 'uniform probability distribution' for SAWs as the distribution which assigns to every element of Ω_0 a constant probability $1/K_0$.

Secondly, an algorithm capable of generating a series of statistically independent and uniformly distributed SAWs must be provided. A naive or 'simple sampling' procedure can be immediately defined [15]. (i) Generate a conventional RW, uniformly distributed in Ω. (ii) Check the RW for self-avoidance. If it is self-avoiding take it as a new sample, otherwise return to (i).

It is not difficult to demonstrate that the series of samples generated by such a process is uniformly distributed in Ω_0. However, it can be shown that in most cases this algorithm is not very efficient: the probability of success in step (ii) is K_0/K, which behaves as $K_0/K \sim \exp(-s/s_0)$, asymptotically with $s \to \infty$, with $s_0 = 2.4, 4.0, 6.0, \ldots, \infty$ for $d = 2, 3, 4, \ldots, \infty$ [15, 23, 24], and this implies that the procedure becomes inefficient for $s \geqslant 10s_0$ approximately.

3.1. The pivot algorithm

The figures of the last paragraph imply that the simple sampling procedure may not be applicable, from the practical point of view, to generate low-dimensional SAWs with s larger than $\simeq$ 50–100. Even if the selection mechanism of step (i) is modified, preconditioning the initial selection in order to enlarge the acceptance probability, as explained in [15], the algorithm continues to be computationally unaffordable when s is large. To overcome this difficulty, a number of improved algorithms have been devised. The detailed description of such procedures is beyond the scope of this work; the interested reader can find them exhaustively reviewed in [15].

Among all the algorithms that have been developed to generate discrete SAWs in low d spaces, the so-called *pivot algorithm* remains one of the most efficient, and is the one we have used for our simulations.

The pivot algorithm [14, 15] is a dynamic procedure, that is, each new SAW is obtained by performing an elemental modification on the previous one. A 'completely different' sample is obtained after a determined number of elemental moves.

Let G be the symmetry group of the lattice, that is, the group of all orthogonal transformations which leave the lattice invariant (i.e. rotations of $\pi/2$ around the axes, axis reflections, etc). Let G_0 be the set of all elements of G excluding the identity. The pivot algorithm can then be defined as follows. (i) A new SAW is obtained from a given previous one performing the following elementary move. Take at random an element g of G_0, and a point k within the SAW ($0 \leqslant k < s$). Apply g to all elements of the walk subsequent to the pivot point $\boldsymbol{r}_k$, using it as the temporary origin; in other words, evaluate

$$\boldsymbol{r}'_l = \boldsymbol{r}_k + g(\boldsymbol{r}_l - \boldsymbol{r}_k) \qquad l = k+1, \ldots, s. \tag{9}$$

(ii) Check the resulting set of points $(\boldsymbol{r}_0, \ldots, \boldsymbol{r}_k, \boldsymbol{r}'_{k+1}, \ldots, \boldsymbol{r}'_s)$ for self-avoidance. If the check is passed then accept the set as the new SAW, otherwise count the previous one once more. This procedure satisfies detailed balance with the correct equilibrium distribution, as well as ergodicity [14, 15, 28]. (iii) In order to obtain statistically independent samples it is necessary to perform a number N_S of elementary moves between samples. The magnitude of N_S varies from case to case, as we shall see below. (iv) It is also necessary to provide as input an initial SAW, randomly selected from the set Ω_0. This can be done either by using another algorithm, or by thermalizing an arbitrary SAW with N_T pivot steps. The first alternative can be used for s not so large; otherwise the second option may be the only applicable one.

A detailed study of the pivot algorithm [14] showed that the acceptance fraction of step (ii) behaves roughly as s^{-p} with s large, where $p \simeq 0.19$ (0.11) for $d = 2$ (3) (for $d > 3$, p goes monotonically to zero). It has also been established that a few $C \sim 10$ accepted moves are generally sufficient to get a practically independent SAW. This leads to the conclusion that taking

$$N_S \geqslant Cs^p \tag{10}$$

will be enough to ensure an adequate statistical independence between samples. We have found that this is acceptable provided that the *minimum* number of accepted moves between samples obtained in a simulation is not much lower than 10. Looking at equation (10) one can see that C actually represents the *average* number of accepted moves between samples. If the dispersion around this mean is important (as it does happen for s very large, and especially for $d = 2$), the mentioned minimum can be very low, even zero, and not completely independent samples can eventually be generated unless a somewhat larger N_S is used. We consider that setting $C \sim 20$ in equation (10) is sufficient to obtain a safe *lower bound* for N_S when s is large (that is, $s \sim 50\,000$). Of course, it is always recommended to choose N_S as large as permitted by the computational power available.

The number of thermalizing steps, N_T, should always be taken large enough to ensure that a stationary regime has been reached. Our experience indicates that $N_T \geqslant 10^3 N_S$ is adequate for most cases.

4. Numerical results

The numerical simulations were performed using the pivot algorithm, already introduced in the previous section, to generate a large number of statistically independent samples, N, keeping d and s fixed. For each sample, the already defined quantities were evaluated and used to perform several statistical analyses. A linear chain of length s was used as the initial configuration. Sampling was started after a thermalizing process of N_T pivot steps. The values used for parameters N_T and N_S were always generously larger than the bounds given in section 3.1.

The frequency histograms for the different observables considered were normalized accordingly with [4]. (i) A certain number m of intervals $[x_k, x_{k+1})$, $k = 1, \ldots, m$ ($x_k > x_l$ for $k > l$) were defined such that $[x_1, x_{m+1}]$ represents the region of interest for the corresponding variable. (ii) The N Monte Carlo samples are used to evaluate the m frequencies f_k, $k = 1, \ldots, m$, which represent the number of times the variable happened to lie within the corresponding interval $[x_k, x_{k+1})$. The f_k are modified to obtain normalized frequencies h_k in the following way:

$$h_k = \frac{f_k}{N(x_{k+1} - x_k)} \qquad k = 1, \ldots, m. \tag{11}$$

The relation between these frequencies and an analytical probability distribution $P(x)$, for N and m large, follows immediately [4]:

$$h_k \leftrightarrow P(\xi_k) \tag{12}$$

where ξ_k can be approximated by

$$\xi_k = \frac{x_{k+1} + x_k}{2}. \tag{13}$$

The actual values of m used in the present simulations vary from case to case, but are always large enough to permit a safe use of equation (13).

4.1. Asymptotic behaviour with $s \to \infty$

A widely studied characteristic of SAWs is the behaviour of some quantities in the limit of large lengths. A typical example is the radius of gyration, whose mean is known to grow with s accordingly with the law [18]

$$\langle S^2 \rangle = M_{S^2} s^{2\nu}(1 + b^{(1)} s^{-\Delta_1} + \cdots) \tag{14}$$

where ν and Δ_1 are referred to as *critical exponents*, and M_{S^2} as *amplitude*. The values of these constants depend on the dimension d. In the case of ν, it is believed that [19–22, 26]

$$\nu = \begin{cases} \frac{3}{4} & \text{for } d = 2 \\ 0.588 & \text{for } d = 3 \\ \frac{1}{2} + \varepsilon/16 + \cdots & \text{for } d = 4 - \varepsilon \\ \frac{1}{2} & \text{for } d > 4. \end{cases} \tag{15}$$

The mean value of any other global observable will have a similar scaling behaviour [22]. This includes the end-to-end distance, r_s^2, the inertia eigenvalues, etc, and all their statistical moments.

In most of the numerical simulations performed to test this scaling behaviour, mean values of 'size' quantities like the radius of gyration were always used. We do not know about any study of the scaling properties of shape quantities like the inertia moments, their probability distributions and/or related observables.

Consider, for instance, the inertia moments λ_k, $k = 1, \ldots, d$. It is expected that, for $s \to \infty$

$$\langle \lambda_k \rangle = M_k s^{2\nu_k}(1 + \cdots) \qquad k = 1, \ldots, d \tag{16}$$

where all the exponents ν_k should be equal to the exponent of the radius of gyration:

$$\nu_k = \nu \qquad k = 1, \ldots, d. \tag{17}$$

It is possible to see that a scaling law of the form of (14), with ν given by (15), also holds for the statistical moments of the inertia eigenvalues, for example, the standard deviations σ_{λ_k}. Some representative examples which confirm such behaviour are presented in table 1, where the mean values of the largest and smallest inertia moments together with the respective standard deviations are tabulated for several values of s, and for $d = 2$ and 3. The moments and the corresponding standard deviations are divided by the leading term of equation (16):

$$\langle \lambda_k \rangle_L = M_k s^{2\nu} \tag{18}$$

with ν given by equation (15). The amplitudes M_k were estimated in the following way: from equations (5), (14) and (18), it is possible to demonstrate that

$$M_k = \frac{\langle \lambda_k \rangle}{\langle S^2 \rangle} M_{S^2}. \tag{19}$$

The amplitude M_{S^2} for the radius of gyration has been extensively studied, and very precise estimates were obtained [22]: $M_{S^2} = 0.108\,15\ (0.194\,55)$ for $d = 2$ (3). Using these data and the data coming from our simulations (for the largest values of s), equation (19) can be evaluated. The resulting amplitudes are presented in table 2.

The data of table 1 reveal that the same scaling law applies to all the mean values and standard deviations considered, as expected. It is also important to note that the observed values are not exactly equal to the asymptotic prediction, and that the difference diminishes when s grows. This indicates that the corrective terms of equation (16) cannot be completely

Table 1. Mean value of the largest and smallest inertia moments ($\langle \lambda_1 \rangle$, $\langle \lambda_d \rangle$) and their respective standard deviations, divided by the asymptotic expectation (18) for the corresponding moment, 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, which correspond to twice the standard error of the mean.

d	s	N ($\times 10^3$)	$\langle \lambda_1 \rangle / \langle \lambda_1 \rangle_L$	$\sigma_{\lambda_1} / \langle \lambda_1 \rangle_L$	$\langle \lambda_d \rangle / \langle \lambda_d \rangle_L$	$\sigma_{\lambda_d} / \langle \lambda_d \rangle_L$
2	100	1000	1.011 07(92)	0.4615	1.011 8(10)	0.5141
	200	1000	1.006 25(93)	0.4639	1.007 5(10)	0.5151
	500	300	1.002 9(17)	0.4639	1.002 9(19)	0.5140
	1 000	450	1.001 5(14)	0.4646	1.003 3(15)	0.5165
	5 000	150	1.000 1(24)	0.4647	1.001 2(27)	0.5139
	10 000	100	1.001 7(30)	0.4688	1.002 3(33)	0.5144
	40 000	50	1.002 7(43)	0.4698	1.003 1(48)	0.5189
3	100	1000	0.962 0(10)	0.5030	0.935 41(77)	0.3840
	200	1000	0.970 8(10)	0.5152	0.952 70(79)	0.3934
	500	300	0.980 3(19)	0.5269	0.970 9(15)	0.4016
	1 000	250	0.986 0(21)	0.5345	0.981 8(16)	0.4068
	5 000	150	0.991 3(28)	0.5404	0.995 3(21)	0.4128
	10 000	100	0.988 6(34)	0.5397	0.996 1(26)	0.4119
	40 000	50	0.995 8(49)	0.5451	0.998 7(37)	0.4140

Table 2. The amplitudes M_k of equation (18), tabulated for $d = 2, 3$.

d	M_1	M_2	M_3
2	0.0935	0.0146	—
3	0.1518	0.0320	0.0108

neglected. Furthermore, the corrections are more significant for $d = 3$, and in this case the asymptotic value is reached *from below* for both the eigenvalue means and the standard deviations, implying that the correction is negative, similarly as with the mean radius of gyration [22].

Let us now consider the probability distribution of the inertia eigenvalues, $P_{k;s}(\lambda)$, where s explicitly indicates the dependence on the length of the walks. If a scaling law similar to that already considered applies to this probability distribution, then a scaling constant $W_{s_2 s_1}$ should exist such that

$$P_{k;s_1}(\lambda) = W_{s_2 s_1} P_{k;s_2}(W_{s_2 s_1}\lambda). \tag{20}$$

It is easy to see that this scaling constant is related to the mean moments via

$$W_{s_2 s_1} = \frac{\langle \lambda_k \rangle_{s_2}}{\langle \lambda_k \rangle_{s_1}} \tag{21}$$

and using equation (16), it can be put in the form

$$W_{s_2 s_1} = \left(\frac{s_2}{s_1}\right)^{2\nu} \left(\frac{1 + \text{higher order corrections } (s_2)}{1 + \text{higher order corrections } (s_1)}\right). \tag{22}$$

We have checked the veracity of equation (20) using our Monte Carlo data. To this end we renormalized the histogram of the scaled distribution $W_{s_2 s_1} P_{k;s_2}(W_{s_2 s_1}\lambda)$ to obtain a modified data set containing the normalized frequencies for the set of m intervals selected for $P_{k;s_1}(\lambda)$ (see equation (11)), and then compared both histograms using an $m \times 2$ chi-square test. The results from such a test are as follows. (i) Using equation (21) to evaluate the scaling constants, for both $d = 2$ and $d = 3$, and for all combinations of k and $s_1 \neq s_2$, the distributions are coincident within the 0.01 significance level†. (ii) The coincidence is maintained if one considers the largest lengths and evaluates $W_{s_2 s_1}$ using the leading term of (22).

Using equations (21) and (22) and the data displayed in table 2, it is easy to see that in the asymptotic regime ($s \to \infty$), the scaling constants $W_{s_2 s_1}$ are independent of k, and that a slight dependence of $W_{s_2 s_1}$ on k appears for lower values of s, especially for $d = 3$. This indicates that the corrections to scaling of equation (22) vary with k, in contrast with the universal nature of the leading term.

A convenient way to study this fact further is to analyse the behaviour of the asphericity (7) when the SAW size goes to infinity. We performed several simulations to check the degree of dependence on the SAW length of the main characteristics of the shape distributions of SAWs. To this end we studied the mean value of the asphericity and its standard deviation. In table 3 we placed our results for s from 100 to 40 000, and in figure 1 the mean asphericity $\langle A \rangle$ is plotted versus s for $d = 2$ (a) and $d = 3$ (b). The error bars represents twice the standard error. From figure 1(a) it remains evident that within error bars the mean value of the two-dimensional asphericity remains constant for all the sizes studied. On the other hand, in the three-dimensional case, plotted in figure 1(b), there is an asymptotic regime, but it is reached only for the largest sizes s, indicating that the corrective terms of equation (16) are more significant in this case. This agrees with previous results obtained for size quantities [22]. We can see that in both cases, the behaviour of the mean asphericity is compatible with both the equality of exponents of equation (17) and the

† To perform a chi-square test adequately it is necessary to discard all the data points with very low frequencies (less than 6). In our case this implies that both 'tails' of the distribution, namely, the region near the origin and the region of large λ, could not be taken into account in the test. Plotting all the checked pairs of distributions we could observe that there is always a very close coincidence in the tails, and for this reason we consider that the conclusions of the chi-square tests can be safely extrapolated to the whole distribution.

Table 3. Mean value and standard deviation of the asphericity ($\langle A\rangle$, σ_A), 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, which correspond to twice the standard error of the mean. The mean values are plotted in figure 1.

d	s	N ($\times 10^3$)	$\langle A\rangle$	σ_A
2	100	1000	0.504 20(50)	0.248 43
	200	1000	0.503 75(50)	0.248 61
	500	300	0.504 28(91)	0.248 21
	1 000	450	0.503 68(74)	0.248 65
	5 000	150	0.503 5(13)	0.248 83
	10 000	100	0.503 7(16)	0.248 85
	40 000	50	0.503 0(22)	0.248 98
3	100	1000	0.438 82(38)	0.191 54
	200	1000	0.436 61(38)	0.191 31
	500	300	0.434 65(70)	0.191 19
	1 000	250	0.433 81(77)	0.191 79
	5 000	150	0.431 40(98)	0.191 23
	10 000	100	0.430 8(12)	0.191 09
	40 000	50	0.430 9(17)	0.191 33

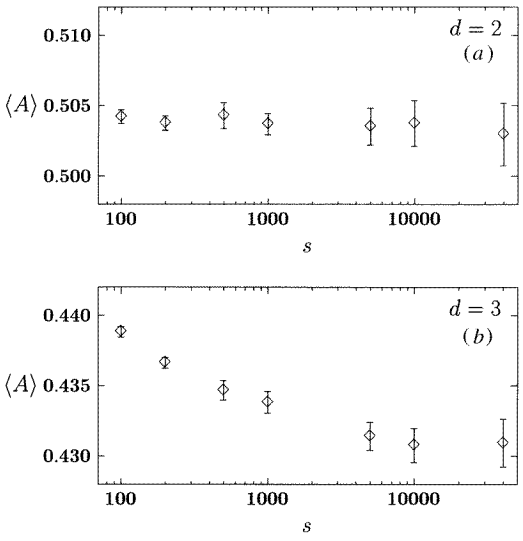

Figure 1. Mean value of the SAW asphericity $\langle A\rangle$ plotted versus s (logarithmic scale) for (a) $d = 2$ and (b) $d = 3$. The error bars represent twice the standard error of the mean. The numerical values are tabulated in table 1. Notice the vertical scales which span very narrow intervals with less than 2% of dispersion around the central value.

already mentioned dependence on k of the scaling factors of equation (22). From the data of table 3 it can also be noticed that the standard deviations remain virtually invariant for both $d = 2, 3$. The behaviour of the mean and standard deviation indicate that the asphericity distribution, $P_A(A)$, $0 \leqslant A \leqslant 1$, should approach a stable asymptotic distribution for large sizes. We have performed chi-square tests, similar to the ones previously mentioned, to compare the distributions for sizes 10 000 and 40 000, respectively, and concluded that for both $d = 2$ and 3 there is coincidence at the 0.01 significance level. The very low frequency regions of the distributions were not included in the tests. These regions, however, do not seem to present a pathological behaviour, and our conclusion is that one can safely extend the results of the chi-square test to the whole range $0 \leqslant A \leqslant 1$. Of course, a more rigorous

proof would require a much larger number of samples.

The data presented so far lead us to the conclusion that for all the sizes s that were considered, that is $100 \leqslant s \leqslant 40\,000$, the global structure of the probability distributions of the different quantities considered in this paper does not present substantial variations, and that this also applies in the asymptotic regime ($s \to \infty$). Therefore, to describe the characteristics of such probability distributions it is sufficient to study them for a single fixed value of s conveniently chosen. We have taken $s = 200$ for most of the simulations of the following sections.

4.2. The principal inertia moments

As we have seen in the previous section, the mean values of the inertia moments $\langle \lambda_k \rangle$ grow as $s^{2\nu}$ as $s \to \infty$ (equations (16) and (17)), with $\nu = \frac{1}{2}$ for RWs, and $\nu > \frac{1}{2}$ (15) for SAWs with $d = 2, 3$. Loosely speaking, this means that for a given length s, and in the average, SAWs should be somewhat 'larger' than conventional RWs.

We also know that the RW probability distributions of the different inertia eigenvalues λ_k can be approximated by chi-square distributions of the form [3]

$$P_k(\lambda) = \frac{1}{\Gamma(\nu_k)} \frac{\nu_k}{\alpha_k} \left(\frac{\nu_k \lambda}{\alpha_k} \right)^{\nu_k - 1} \exp\left(-\frac{\nu_k \lambda}{\alpha_k} \right). \tag{23}$$

$\Gamma(x)$ stands for the gamma function, and α_k and ν_k are external parameters.

For the case of SAWs, one expects that the excluded volume effect will accordingly modify the inertia eigenvalue probability distributions. We have studied these distributions, and effectively found out that they are different in the two cases. In figure 2(a) we show the probability distribution of λ_1 in the two-dimensional case. We have plotted the RW

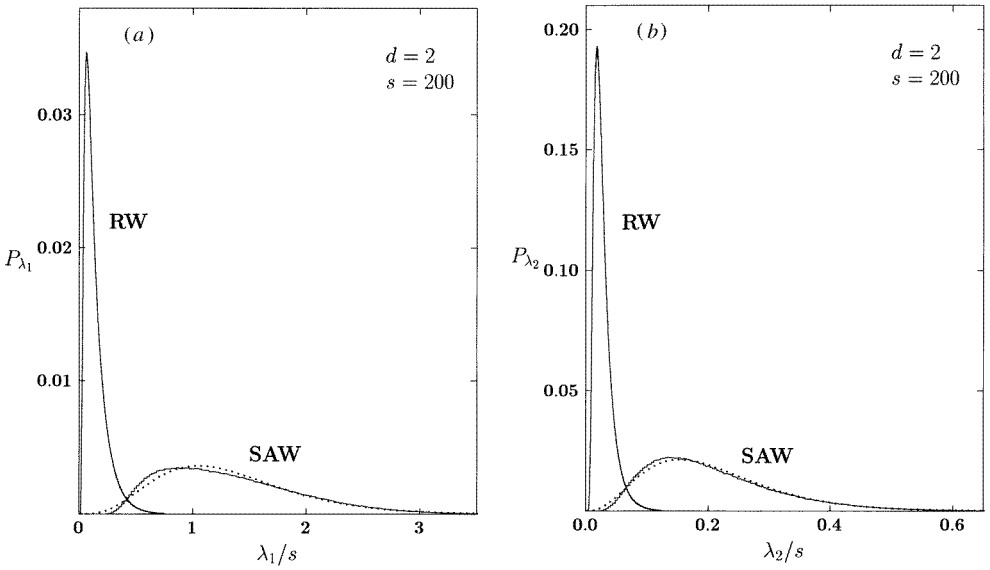

Figure 2. Probability distributions of the inertia moments (a) λ_1 and (b) λ_2, for $d = 2$ and $s = 200$. The normalized histograms in full curves represent the Monte Carlo data for SAWs and RWs with $N = 10^6$ and 2×10^7, respectively. The dotted curve represents distribution (23). For convenience, a scaling factor s^{-1} was used in the horizontal axis.

distribution (obtained with an independent simulation employing the method of [3], and using $N = 2 \times 10^7$ samples), as well as the SAW distribution. The number of SAW samples used is $N = 10^6$. The difference between both distributions is evident. First we notice the expected shift of the SAW distribution to higher values of λ_1. The second difference is that the SAW distribution spreads around its most probable value much more than in the RW case, that is, the standard deviation of the SAW distribution is significantly larger than the corresponding one for the distribution of the common RWs. This, of course, is also related to the scaling laws of the distributions and their statistical moments, discussed in section 4.1.

As mentioned before, the probability distribution for the eigenvalues λ_k can be

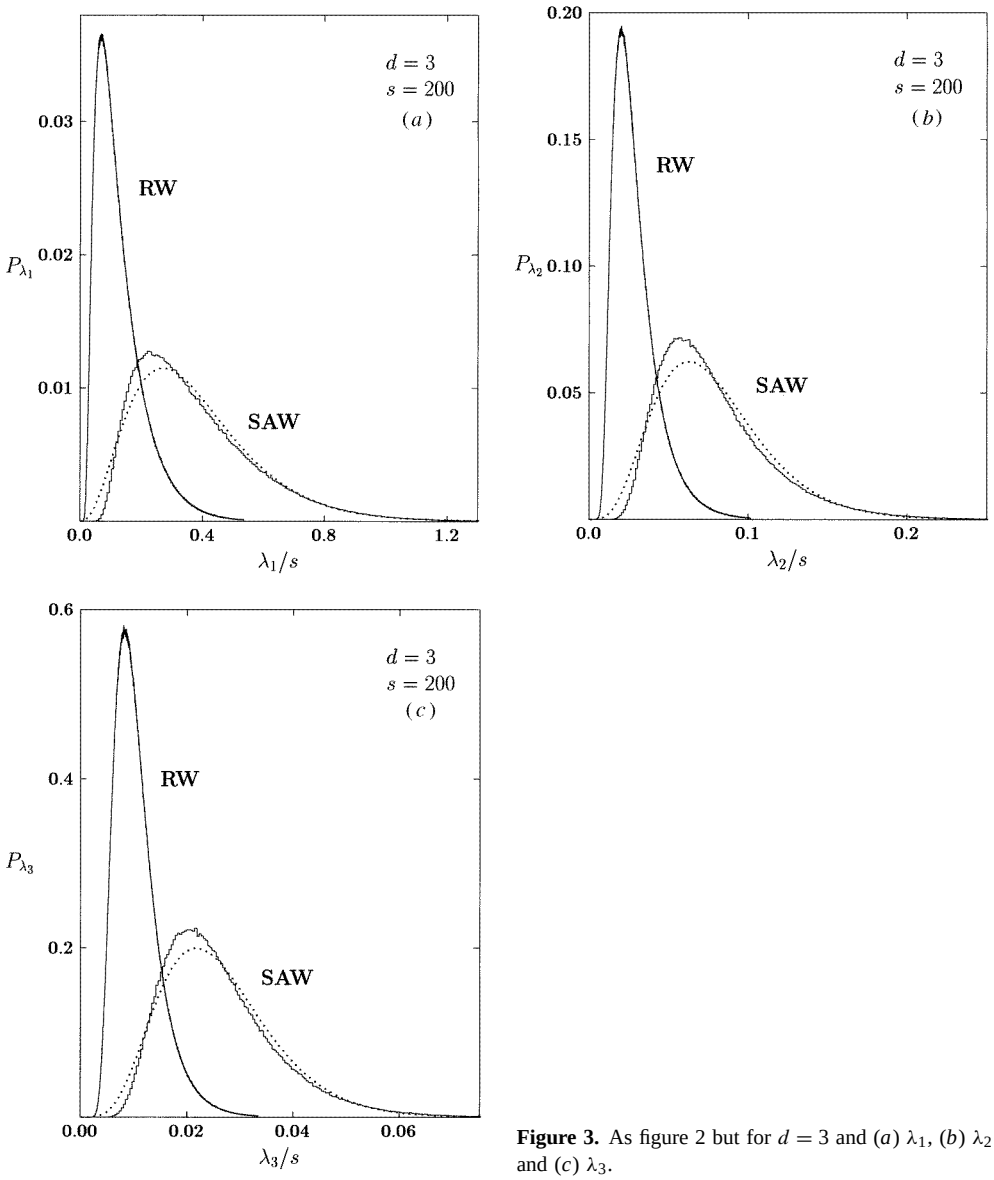

Figure 3. As figure 2 but for $d = 3$ and (*a*) λ_1, (*b*) λ_2 and (*c*) λ_3.

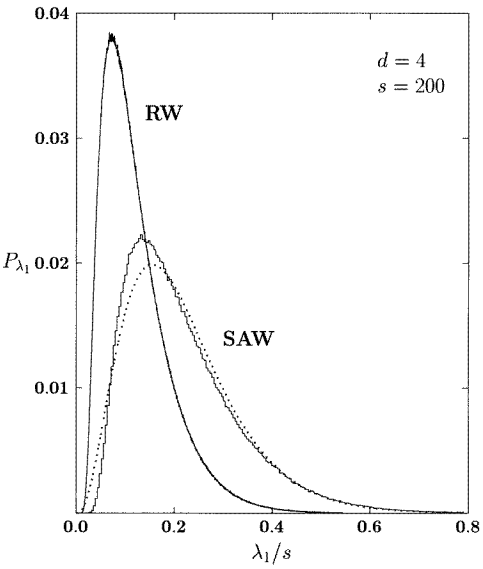

Figure 4. As figure 2 but for λ_1 in the $d = 4$ case.

approximated, in the case of unrestricted RWs, by a chi-square distribution of the form (23). For simplicity, this approximate distribution was not plotted in figure 2(*a*). We were interested in checking whether the same distribution would fit adequately in the SAW case, and therefore we used the method already explained in [3] to evaluate the distribution parameters. The result is plotted in figure 2(*a*) (dotted curve) and we can see that both distributions are not strictly coincident but there exists an acceptable agreement, albeit not so good as in the RW case [3].

Qualitatively similar results are presented in figure 2(*b*) for the case of λ_2. In this case we can observe the different scale in the x-axis stressing the fact that the second inertia eigenvalue is on average significantly smaller than λ_1. The differences between the RW and SAW cases remain the same as in the λ_1 case: the SAW probability distribution for λ_2 is located in a shifted position with respect to the RW distribution, and possesses a larger dispersion around its most probable value.

We have also studied the eigenvalue probability distributions in spaces of higher dimension. The results for $d = 3$ are placed in figure 3. When comparing the SAW and RW distributions, we can observe that the first ones present a shift in the most probable value location, and wider distributions, similarly as reported for the $d = 2$ case. In this case, however, the shift is less significant. This tendency is maintained when d is increased, as can be seen in figure 4, which displays the four-dimensional distribution of λ_1. In this case the most probable values are not much different, but the SAW dispersion is longer than in the RW case. For $d \to \infty$ a SAW and an RW become equivalent objects (as measured as their inertia eigenvalues) [27], and so both distributions should become identical in such a limit. Note that the probability distributions for the SAW inertia eigenvalues differ from the respective RW ones even when the space dimension is greater than or equal to the critical dimension for SAWs ($d \geqslant d_c = 4$).

4.3. Inertia moment ratios

The ratios R_{ij} are a set of observables traditionally used to study the shape of RWs in low d spaces. In a previous work [4] we made a detailed study of the corresponding probability

distributions $P_{R_{ij}}$, arriving at the conclusion that the distribution

$$P_{R_{ij}}(z) = \begin{cases} \dfrac{1}{U}\dfrac{z^{\omega-1}(1-z)^{\gamma-1}}{(1+az)^{\omega+\omega'}} & 0 \leqslant z \leqslant 1 \\ 0 & z > 1 \end{cases} \tag{24}$$

fits excellently with the true ones in all cases. ω, ω', γ and a are external parameters, and U is a normalizing constant given by [4]

$$U = \frac{\Gamma(\gamma)\Gamma(\omega)}{\Gamma(\gamma+\omega)}(1+a)^{-\omega}F\left(\gamma-\omega', \omega; \gamma+\omega; \frac{a}{1+a}\right) \tag{25}$$

where the function $F(\alpha, \beta; \gamma; z)$ is the well known hypergeometric function.

The parameters ω, ω', γ and a vary from case to case. They must verify the following constraints [4]:

$$\omega > 1 \qquad \omega' \geqslant 0 \qquad \gamma > 1 \qquad a > 0 \tag{26}$$

and are related to the parameters α and ν of distribution (23) in the limit $d \to \infty$, $i \gg j$ [4]:

$$\begin{aligned} &\omega \to \nu_i \\ &\omega' \to \nu_j \\ &\gamma \to 1 \\ &a \to (\alpha_j \nu_i)/(\alpha_i \nu_j). \end{aligned} \tag{27}$$

For low d, however, the parameters must be evaluated numerically.

The fact that the probability distribution of the low d SAW inertia eigenvalues spreads around the most probable value much more widely than the corresponding ones for common RWs might lead to the conclusion that the same would happen with the probability distribution of the inertia moment ratios R_{ij}. The results of our simulations indicate that this conclusion is not true. In fact, the probability distributions of SAW and RW moment ratios are, in general, very similar, the SAW distributions being slightly more peaked than the RW ones, indicating clearly that the probability distributions for the different inertia eigenvalues are not independent: the independence between the distributions $P_i(\lambda_i)$ and $P_j(\lambda_j)$ would imply a different $P_{R_{ij}}$ (see reference [4]).

In figures 5 and 6 we display the results obtained in representative cases: figure 5 corresponds to $P_{R_{21}}$, $d = 2$; and figure 6 corresponds to the distributions in three-dimensional spaces $P_{R_{21}}$ (a), $P_{R_{31}}$ (b), and $P_{R_{32}}$ (c). The dotted curves correspond to the RW distribution while the full histograms corresponds to the SAW case. $N = 10^6$ (2×10^7) for SAWs (RWs). It is evident from all these plots that both SAW and RW distributions are not much different, and that the existing difference between distributions is less significant for $d = 3$. For $d \geqslant 4$ both distributions become virtually coincident.

An additional plot is present in figures 5 and 6: the full curves, not very distinguishable from the corresponding SAW histograms, represent a nonlinear least-squares fit of distribution (24) to the Monte Carlo data. As in the common RW case [4], the agreement between the Monte Carlo data and the analytical distribution is excellent, and allows us to conclude that the mentioned distribution can be taken as a very good approximation to the true one for both RW and SAW cases, and for all ratios R_{ij} $(i > j)$. In table 4 we give the values taken by the parameters of distribution (24) obtained from the nonlinear fits.

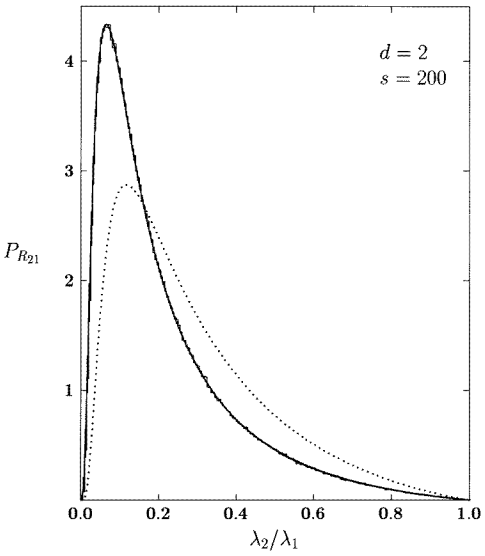

Figure 5. Probability distributions of the inertia moment ratio R_{21} for $d = 2$ and $s = 200$. The normalized histogram in full curve represents the Monte Carlo data for SAWs with $N = 10^6$ samples, while the dotted curve corresponds to RWs with $N = 2 \times 10^7$ samples. The full continuous curve (not very distinguishable from the SAW histogram) corresponds to distribution (24) with parameters obtained from a nonlinear least-squares fit to the Monte Carlo data [4].

Table 4. Representative values of parameters corresponding to the probability distribution (24) for different ratios R_{ij}, as obtained from nonlinear least-squares fits of (24) to the respective Monte Carlo data [4].

d	i	j	ω	ω'	γ	a
2	2	1	6.40	0.58	2.02	49.37
3	2	1	8.54	0.79	1.99	31.65
3	3	1	8.73	1.11	6.00	77.11
3	3	2	7.92	0.78	2.08	11.73

4.4. Asphericity and angle Θ

Another interesting observable to measure in a Monte Carlo simulation is the asphericity (7).

For $d = 2$ this observable is related to the ratio R_{21} via [4]

$$A = \left(\frac{1 - R_{21}}{1 + R_{21}}\right)^2. \tag{28}$$

This means that the probability distribution of the asphericity, P_A, is connected to $P_{R_{21}}$. In fact, from [4] and using the fact that $P_{R_{21}}(z) = 0$ for $z \geqslant 1$, we obtain

$$P_A(x) = \frac{1}{\sqrt{x}(1 + \sqrt{x})^2} P_{R_{21}}(\beta(x)) \tag{29}$$

where $0 \leqslant x \leqslant 1$ and

$$\beta(x) = \frac{1 - \sqrt{x}}{1 + \sqrt{x}}. \tag{30}$$

If $P_{R_{21}}$ is given by (24), then P_A can be written as follows:

$$P_A(x) = \frac{2^{\gamma-1}}{U} \frac{(\sqrt{x})^{\gamma-2}(1 - \sqrt{x})^{\omega-1}}{(1 + \sqrt{x})^{2-\omega'}[(a+1) - (a-1)\sqrt{x}]^{\omega+\omega'}}. \tag{31}$$

In figure 7 the results of our simulations for $d = 2$ are displayed. The full (dotted) histogram corresponds to the SAW (RW) case, with N equal to 10^6 (2×10^7). Again, like in

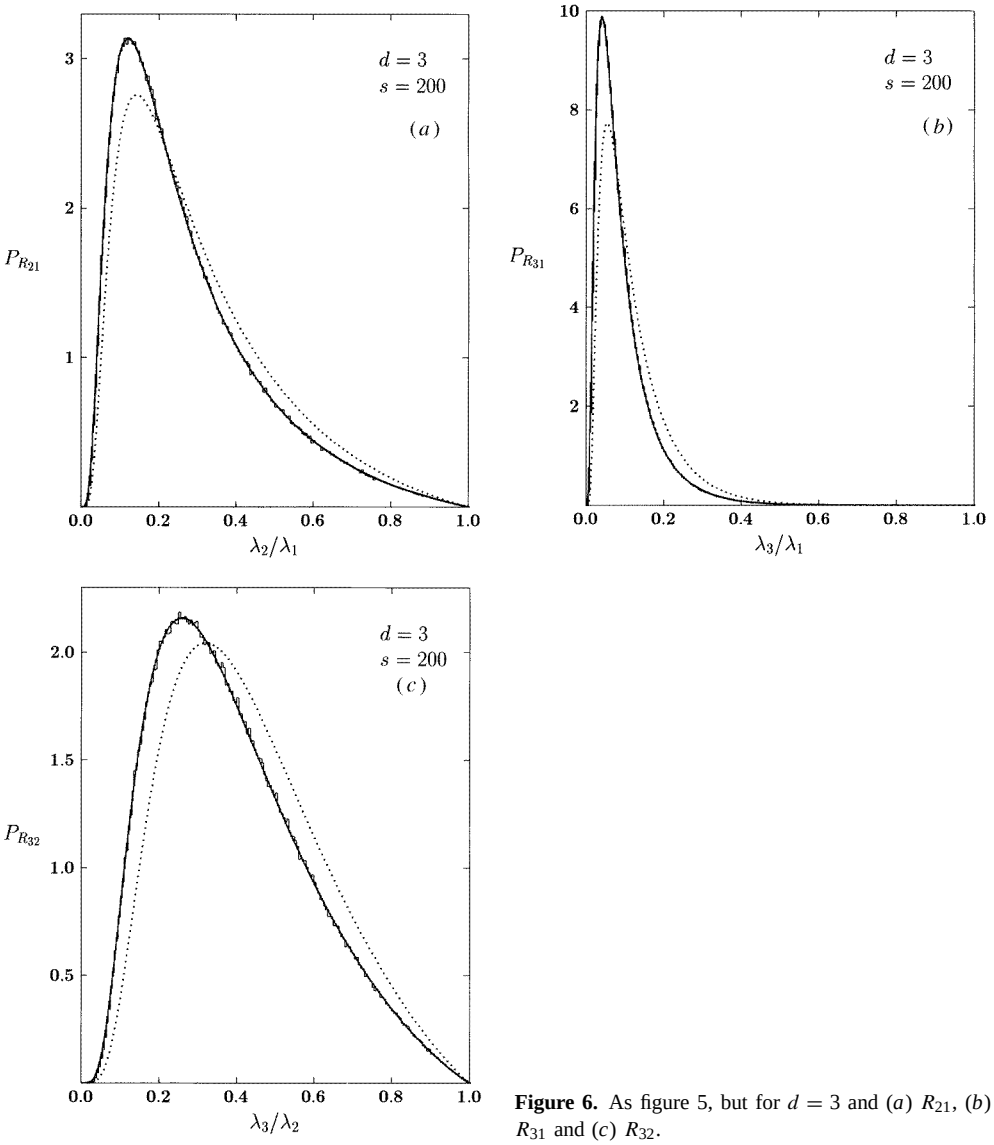

Figure 6. As figure 5, but for $d = 3$ and (*a*) R_{21}, (*b*) R_{31} and (*c*) R_{32}.

the case of the moment ratios, the correlations between the inertia eigenvalue probabilities show up in a somewhat sharp distribution for the asphericity, which presents a peak for $A \simeq 0.7$. For $A \to 0$ the SAW distribution is significantly smaller than the RW one, but it is always non-zero, even at the origin.

The full continuous curve which goes along the SAW histogram corresponds to distribution (31), using the parameters of table 4 for R_{21} and $d = 2$; no fit was performed using the asphericity simulation data.

It is evident that there exists an excellent concordance between distribution (31) and the true one. Notice that the fact that $P_A(0) \neq 0$ requires that $\gamma = 2$ [4]. The value $\gamma = 2.02$ obtained from the $P_{R_{21}}$ fit (table 4) is consistent with this requirement, and the small difference cannot be considered significant. Replacing $\gamma = 2$ in equation (31) we

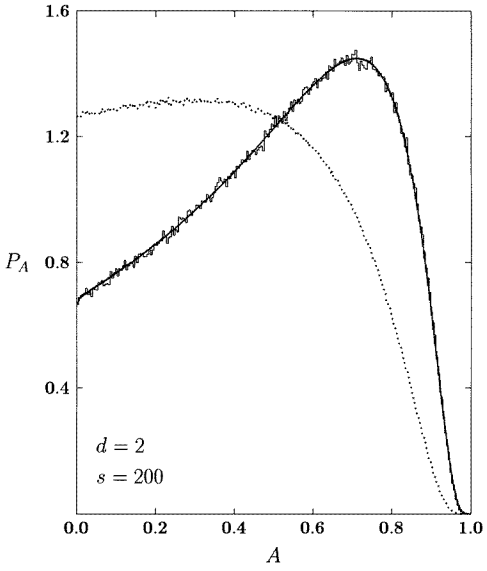

Figure 7. Probability distributions of the asphericity A for $d = 2$ and $s = 200$. The normalized histogram in full (dotted) curve represents the Monte Carlo data for SAWs (RWs) with $N = 10^6$ (2×10^7) samples. The full continuous curve (not very distinguishable from the SAW histogram) corresponds to distribution (31) with parameters taken from table 2 for the corresponding $d = 2$ case (no fit was performed using the asphericity Monte Carlo data).

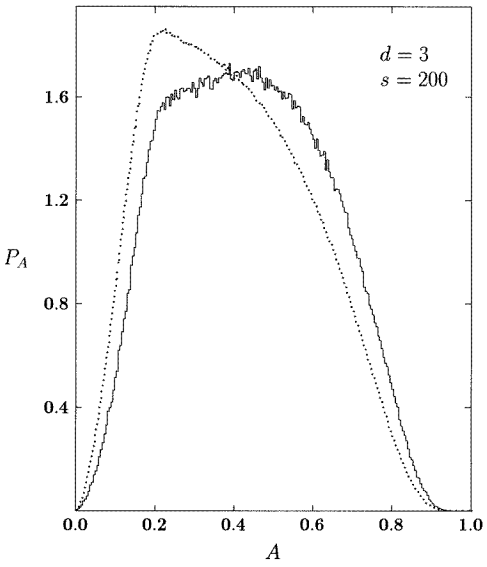

Figure 8. Probability distributions of the asphericity A for $d = 3$ and $s = 200$. The normalized histogram in full (dotted) curve represents the Monte Carlo data for SAWs (RWs) with $N = 10^6$ (2×10^7) samples.

obtain

$$P_A(x) = \frac{2}{U} \frac{(1-\sqrt{x})^{\omega-1}}{(1+\sqrt{x})^{2-\omega'}[(a+1)-(a-1)\sqrt{x}]^{\omega+\omega'}} \tag{32}$$

as an analytical distribution of the asphericity in two dimensions.

We have also studied the asphericity in spaces of higher dimensions. Figure 8 shows our results for $d = 3$. The full (dotted) histogram corresponds to the SAW (RW) case, with $N = 10^6$ (2×10^7). It can be seen that both distributions do not present important differences. For $d > 3$ the differences between curves becomes even smaller, and in fact both curves are virtually coincident for $d \geqslant 4$.

We also want to mention that a previous simulation of the three-dimensional

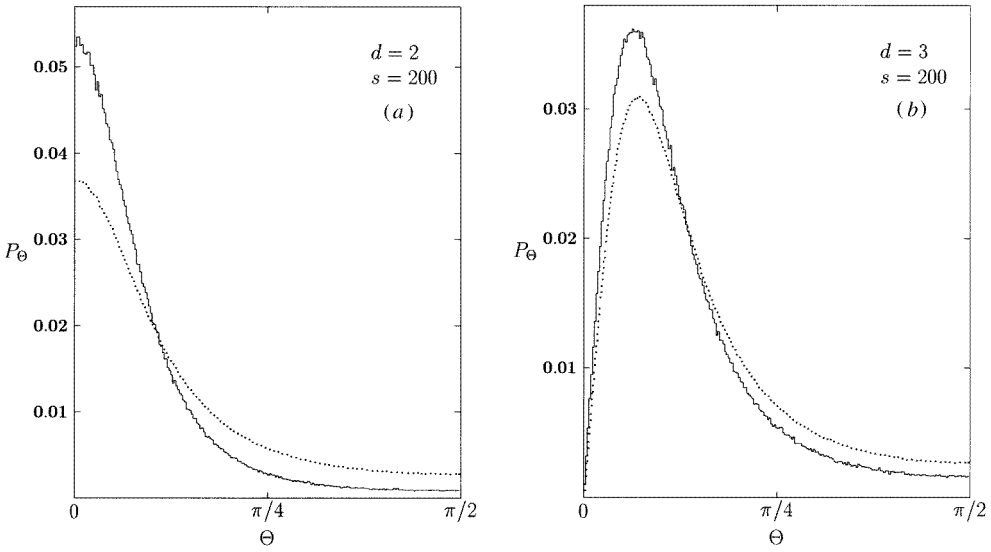

Figure 9. Probability distributions of the angle Θ (equation (8)) for $s = 200$ and (a) $d = 2$, (b) $d = 3$. The normalized histogram in full (dotted) curve represents the Monte Carlo data for SAWs (RWs) with $N = 10^6$ (2×10^7) samples.

asphericity [8] for $s = 140$ is in acceptable agreement with the present one. We recall, however, that in the simulation of [8], the number of samples used (10^5) is one order of magnitude smaller than in our simulation.

We conclude this section presenting the simulation data for the angle Θ defined in equation (8). A non-uniform distribution for Θ indicates that there exist correlations between the end-to-end vector $\boldsymbol{r}_s$ and the principal axis of inertia $\boldsymbol{u}_1$ [3].

In figure 9 we present the distributions P_Θ, in the $d = 2$ (a) and $d = 3$ (b) cases, and for SAWs (full histograms) and RWs (dotted histograms). The number of samples, N, is 10^6 (2×10^7) for SAWs (RWs).

We can see that in all cases both distributions are similar, the SAW curve being slightly more peaked than the RW one. A remarkable fact is that the most probable value is virtually the same for both SAWs and RWs, suggesting that this quantity does not depend upon the self-avoidance condition (2). Notice that for both SAWs and RWs, $P_\Theta(0) \neq 0$ ($= 0$) for $d = 2$ ($d > 2$) [3]. For $d > 3$ and like in the cases of the inertia moment ratios and the asphericity, both SAW and RW probability distributions become practically coincident.

In [9] the mean value $\langle \cos \Theta \rangle$ was calculated for SAWs of different lengths s, ranging from 8 to 1024. We have compared our data with such calculations finding no discrepancies (the number of samples N used in [9] is less than or equal to 10^5).

5. Conclusions and final remarks

We have exhaustively analysed the shape properties of self-avoiding walks (SAWs). The probability distributions of the principal inertia moments, λ_k, their ratios, R_{ij}, the asphericity, A, and the first quadrant angle between the end-to-end-vector and the principal axis of inertia, Θ, were evaluated in a variety of cases, together with related statistical parameters.

The Monte Carlo simulations were performed keeping the space dimension, d, and

the size of the walks, s, fixed. The pivot algorithm, described in section 3.1, was used to generate the SAW samples in all cases. The probability distributions were estimated numerically by means of adequately normalized frequency histograms. The number of independent samples used in most cases, typically 10^6 for SAWs and 2×10^7 for conventional random walks (RWs), is large enough to obtain a high-precision representation of the distributions.

The dependence of these distributions on the SAW size was first considered, arriving at the following conclusions. (i) The distributions for the inertia eigenvalues obey the same scaling laws as the traditional global observables (i.e. the radius of gyration, the squared end-to-end distance, etc) (equation (14)). (ii) The corrections to scaling for the eigenvalues maintain the sign of the corrections for the radius of gyration, reported in [22], that is, positive (negative) for two (three) dimensions. These corrections are not equal for every eigenvalue, especially in the three-dimensional case. (iii) The probability distributions of shape-describing quantities like the asphericity present only small variations as the size of the walks tends to infinity. These small variations are more significant in three than in two dimensions.

We have used our Monte Carlo data to obtain estimates of the inertia moment amplitudes (19), placed in table 3.

Then we have performed an exhaustive analysis of several shape-describing observables for fixed walk length $s = 200$. Both SAW and conventional RW distributions were calculated at each case in order to compare the shape properties of both objects.

The SAW inertia eigenvalue distributions are always wider than the respective RW ones, and are of course shifted towards larger sizes due to the already mentioned scaling properties (see section 4.1). This means that, on average, a SAW is 'bigger' than an RW of equal length, but its size is less predictable. Furthermore, the existence of overlaps between distributions (see figures 2 and 3) indicates that a significant fraction of the SAWs that can be generated in a Monte Carlo simulation will occupy a similar, or even smaller, volume (or area) than the average RWs.

On the other hand, the analysis of the other quantities, like the ratios of inertia moments, the asphericity or the angle between the principal inertia axis and the end-to-end vector, shows us that the corresponding probability distributions are somewhat more peaked than the respective RW ones. This means that SAWs possess a more definite *shape*, in contrast to a more determined *size* for the RWs, as stated in the previous paragraph.

The SAW probability distributions for the ratios R_{ij} were used to check the validity of the analytic distribution (24), which was used in [4] for the case of conventional RWs. In all the cases considered the agreement is excellent, and this lead us to the conclusion that the mentioned distribution (24) is a very good approximation of the true ones for both RWs and SAWs.

The analytic distribution for the two-dimensional asphericity (31) (obtained from the distribution for the ratio R_{21}, as explained in section 4.4), evaluated using the parameters coming from the fit to the two-dimensional R_{21} distribution (table 4), has proved to be practically coincident with the true distribution (see figure 7), obtained from an independent Monte Carlo simulation.

These two last results are consistent with the conclusion that the functional form of the probability distributions of both the ratios R_{ij} and the two-dimensional asphericity seem to be of universal nature, not dependent on the 'interactions' which differentiate SAWs from RWs. This universality refers only to the functional form of the distributions, since the values of the parameters of equations (24) and (32) do vary when passing from RWs to SAWs. Note that the asphericity in d dimensions can be put as a function of the ratios

$R_{21}, R_{31}, \ldots, R_{d1}$ (see equations (6) and (7)), and since the mentioned properties of the ratios apply for the arbitrary dimension case, the conclusions for the asphericity should also be true for all spatial dimensions. We leave it for a future work to analyse the origin and implications of these conclusions in a more complete and rigorous way.

Acknowledgments

The author is indebted to J L Alessandrini for helpful comments and suggestions, and is grateful to Fundación Antorchas of Argentina for its help and financial support. This work was partially supported by the Consejo Nacional de Investigaciones Científicas y Técnicas of Argentina.

References

[1] Fernandez R, Frölich J and Sokal A D 1992 *Random Walks, Critical Phenomena and Triviality in Quantum Field Theory* (Berlin: Springer)
[2] de Gennes P G 1979 *Scaling Concepts in Polymer Physics* (Ithaca, NY: Cornell University Press)
[3] Sciutto S J 1994 *J. Phys. A: Math. Gen.* **27** 7015
[4] Sciutto S J 1995 *J. Phys. A: Math. Gen.* **28** 3667
[5] Kranbuehl D E and Verdier P H *J. Chem. Phys.* **67** 361
[6] Aronovitz J A and Nelson D R 1986 *J. Physique* **47** 1445
[7] Gaspari G, Rudnick J and Beldjenna A 1987 *J. Phys. A: Math. Gen.* **20** 3393
[8] Jagodzinski O, Eisenriegler E and Kremer K 1992 *J. Physique* I **2** 2243
[9] Bruns W 1992 *Makromol. Chem. Theory Simul.* **1** 287
[10] Gaspari G, Rudnick J and Beldjenna A 1993 *J. Phys. A: Math. Gen.* **26** 1
[11] Bird R B 1983 *Physica* **118A** 3
[12] Szleifer I 1992 *J. Chem. Phys.* **92** 6940
[13] Alessandrini J L and Vila J 1994 *Phys. Rev.* E **49** R3584
[14] Madras N and Sokal A D 1988 *J. Stat. Phys.* **50** 109
[15] Sokal A D 1995 *Monte Carlo and Molecular Dynamics Simulations in Polymer Science* ed K Binder (New York: Oxford University Press)
[16] Freed K F 1987 *Renormalization Group Theory of Macromolecules* (New York: Wiley)
[17] des Cloizeaux J and Jannink G 1990 *Polymers in Solution: Their Modelling and Structure* (New York: Oxford University Press)
[18] Nickel B G 1991 *Macromolecules* **24** 1358
[19] Muthukumar M and Nickel B G 1987 *J. Chem. Phys.* **86** 460
[20] Le Guillou J C and Zinn-Justin J 1989 *J. Physique* **50** 1365
[21] Nienhuis B 1984 *J. Stat. Phys.* **34** 731
[22] Li B, Madras N and Sokal A D 1994 Critical exponents, hyperscaling and universal amplitude ratios for two- and three-dimensional self-avoiding walks *Preprint* New York University (NYU-TH-94/09/01); also available from the Los Alamos e-Print server (Website http://xxx.lanl.gov/), preprint number HEP-LAT-9409003
[23] Madras N and Slade G *The Self-Avoiding Walk* (Boston, MA: Birkhäuser)
[24] Hara T, Slade G and Sokal A D 1993 *J. Stat. Phys.* **72** 479
[25] Madras N, Orlitsky A and Shepp L A 1990 *J. Stat. Phys.* **58** 159
[26] Brézin E, Le Guillou J C and Zinn-Justin J 1976 *Phase Transitions and Critical Phenomena* vol 6, ed C Domb and M S Green (London: Academic)
[27] Rudnick J, Beldjenna A and Gaspari G 1987 *J. Phys. A: Math. Gen.* **20** 971