Some enumeration theorems for self-avoiding walks I.

Daisy Joseph & A S Padmanabhan*
School of Chemical Sciences, Mahatma Gandhi University
Priyadarshini Hills P.O., Kottayam 686 560, India
e-mail: padman58@md4.vsnl.net.in

Received 29 September 1999; accepted 15 November 1999

Suppose \((\delta(n))^2\) represents the mean square end to end distance of simple random walks of 'n' steps in D-dimensional Euclidean space it is shown that for simple random walks with exclusion of immediate reversals in D-dimensional Euclidean space:

\[
(\delta(n))^2 = \frac{Dn}{D-1} \sum_{i=0}^{1} \sum_{j=0}^{D-i-1} \frac{1}{2(D-1)^2} \left[1 - \frac{1}{(2D-1)^2}\right]
\]

For random walks with exclusion of immediate reversals and for self-avoiding walks in D-dimensional Euclidean space it is proved that the following holds:

\[
(\delta(n))^2 = \sum_{i=0}^{n-1} \sum_{j=0}^{D} A_i \prod_{j=0}^{((2D)-2j)}
\]

where \(A_i\) and \(R_j\) are independent of D.

For random walks with exclusion of immediate reversals it is shown that for \(A_{n,i}\) (where \(A_{n,i}\) represents \(A_{n,1}\) for 'n' steps) the following recursion formulae hold:

\[
A_{n+1,k} = A_{n,k-1} + (2k-1)A_{n,k} \quad \text{and} \quad A_{n+1,k} = \sum_{j=0}^{2k+1} C_j A_{n,k+1} + A_{n,k}
\]

For self-avoiding walks expressions are derived for \(A_{n,i}, R_{n,i}, A_{n,2}, R_{n,2}, A_{n,3}, R_{n,3}\) and \(A_{n,4}\). Values of \(A_i\) have been reported for all \(i\) for 'n' up to 'n' = 14 and \(R_i\) for all \(i\) for 'n' up to 'n' = 14. Finally it is proved that if \(q_{2n}\) is the number of self-avoiding polygons of 2n steps in D-dimensional Euclidean space then

\[
q_{2n} = (2D) \sum_{i=1}^{n} P_i \prod_{j=1}^{((2D)-2j)} n \geq 2
\]

where \(P_i\) are independent of D. Values of \(P_i\) for all \(i\) for 2n up to 2n = 14 have also been reported.

Finally a new problem has been described in Graph theory and how this problem could be studied further has been outlined. Using the definition of the problem three different types of computer algorithms for Theorems 2, 4 and 5 have been defined.

Introduction

The problem of the self-avoiding walk had its origins in chemical physics as a model for the configurational statistics of linear chain molecules in dilute solution. Since the pioneering work of Flory many attempts have been made to solve exactly the problem of the number of self-avoiding walks of 'n' steps and also the problem of the mean square end-to-end distance of self-avoiding walks of 'n' steps. The latter is especially important in developing a theoretical understanding of experimental
data, such as viscosity of dilute polymer solutions and also light-scattering profile of dilute polymer solutions.

Despite the commendable work of many authors for the past five decades exact formulae for the number of self-avoiding walks, the mean square end-to-end distance in terms of the dimensionality D and the number of steps \( n \), have not been forthcoming. In addition to the search for an exact formula Guttmann et al.\(^{2,30}\) and Sykes et al.\(^{21,22}\) have attempted to develop computer algorithms that enumerate self-avoiding walks. In this area too the ideal of developing a polynomial time algorithm or else proving that the problem is non-deterministic polynomial complete (N. P. complete) seems to be evasive.\(^{23,24}\) The reader is referred to the book by Madras and Slade\(^{25}\) for a better understanding of the efforts that have been made to solve this problem.

In this paper the authors first derive an exact formula for the mean square end-to-end distance \( \langle \sigma_2(n) \rangle \) for simple random walks of \( n \)-steps and D dimension with exclusion of immediate reversals. The authors then express (in terms of a theorem which they prove) \( \langle \sigma_0(n) \rangle \) for self-avoiding walks as a ratio of two power series in D. In the next theorem a similar result is shown to hold for random walks with exclusion of immediate reversals and simple triangular and vertical recursion formulae for the constants derived therein. In the next theorem some, though not all, the constants in the two power series are derived. In the fourth theorem it is shown that a similar result holds for self-avoiding polygons too. The authors then describe a new problem in Graph theory and outline how this problem could be studied further. Using the definition of the problem they then define three different types of computer algorithms for Theorems 2, 4 and 5. Tables that enumerate all constants for \( \langle \sigma_2(n) \rangle \) for \( n \) up to \( n=14 \) and for self avoiding polygons for \( 2n \) up to \( 2n=14 \) are given and comparison with computer enumeration data is discussed in conclusion.\(^{2,20,25}\)

**Theory and Notation**

Consider the set \( \mathbb{Z}^D \), the subset of D-dimensional Euclidean space consisting of those points with integer coordinates. The distance between two points \( \vec{a} = (a_1, a_2, \ldots, a_D) \) and \( \vec{b} = (b_1, b_2, \ldots, b_D) \) in \( \mathbb{Z}^D \) is defined as usual by

\[
|\vec{a} - \vec{b}| = \left( \sum_{i=1}^{D} (a_i - b_i)^2 \right)^{1/2}
\]

We shall be concerned with walks in \( \mathbb{Z}^D \), defined by sequences of displacements (or ‘steps’) \( \{\vec{X}_j\}_{j=1}^{n} \) in \( \mathbb{Z}^D \), with the property that for each \( i \), only one component of \( \vec{X}_i \) is non zero, that component taking either of the values \( \pm 1 \), so that \( |\vec{X}_i| = 1 \) walks commence at the origin, so that the position after \( n \) steps is given by

\[
\vec{S}_n = \sum_{i=1}^{n} \vec{X}_i \quad \text{with} \quad \vec{S}_0 = (0, 0, \ldots, 0)
\]

Let \( \langle \sigma_0(n) \rangle = \langle \vec{S}_n \cdot \vec{S}_n \rangle \) denote the mean square end-to-end distance in an \( n \)-step walk.

We consider two kinds of walks. An \( n \)-step walk with immediate reversals excluded is defined by requiring that \( \vec{X}_{i+1} \neq -\vec{X}_i \) for \( i = 1, 2, \ldots, n-1 \). For such a walk, the 2\( D-1 \) allowed displacements corresponding to each step are assigned equal probabilities \( (2D-1)^{-1} \). An \( n \)-step self-avoiding walk (SAW) is defined by the requirement that \( \vec{S}_k \neq \vec{S}_m \) unless \( k = m \) for \( 0 \leq k \leq n \) and \( 0 \leq m \leq n \). If there are \( C_n \) such \( n \)-step walks, each walk is assigned probability \( 1/C_n \) in calculating the mean-square displacement. Finally, an \( n \)-sided self-avoiding polygon has \( \vec{S}_n = \vec{S}_0 \) but \( \vec{S}_k \neq \vec{S}_m \) unless \( k = m \) for \( 0 \leq k \leq n-1 \) and \( 0 \leq m \leq n-1 \).

**Theorem 1**

In dimension \( D > 1 \), for an \( n \)-step walk with immediate reversals excluded,

\[
\langle \sigma_0(n) \rangle = \frac{D^n}{D-1} - \frac{1}{D-1} - \frac{1}{2(D-1)^2} \left( \frac{1}{(2D-1)^{n+1}} - \frac{1}{(2D-1)^{n+2}} \right)
\]

**Proof**

For \( n \geq 2 \) we have

\[
\langle \sigma_0(n) \rangle = \sum_{i=1}^{n} \langle \vec{X}_i \rangle^2 = \sum_{i=1}^{n} \sum_{j=1}^{n} \langle \vec{X}_i \cdot \vec{X}_j \rangle
\]

\[
= \sum_{i=1}^{n} |\vec{X}_i|^2 + \sum_{j=2}^{n} \sum_{i=1}^{n-j} \langle \vec{X}_i \cdot \vec{X}_j \rangle
\]
n = 2 \sum_{i=2}^{n} \sum_{j=1}^{i-1} (2D-2)^{(i-j)}

Here we have noted that \(|\tilde{X}_i| = 1\). Also, \(\langle \tilde{X}_i \tilde{X}_j \rangle = (2D-1)^{\delta_{ij}}\) for \(j > i\), since only cases in which \(\tilde{X}_j = \pm \tilde{X}_i\) contribute to the average. If there is a bend in the walk, the cases \(\tilde{X}_j = \tilde{X}_i\), and \(\tilde{X}_j = -\tilde{X}_i\) occur with equal frequency and their contributions cancel, so only walks with \(j-i\) steps in the same direction affect the average. The sums over \(i\) and \(j\) are finite geometric progressions, and we find after a little algebra that the stated formula holds for \(n \geq 2\). The formula also reduces to the trivial result \(\langle \tilde{X}_0^2 \rangle = 1\) for \(n = 1\).

**Theorem 2**

\[ \langle \tilde{X}_0^2 \rangle = \sum_{i=0}^{n} \sum_{j=0}^{D-2} R_i \prod_{i=0}^{j} (2D-2j) \]

where \(A_i\) and \(R_i\) are independent of \(D\) and depend only on \(n\) and \(i\).

**Proof**

The basic idea is to partition walks according to the number of axis directions in which steps are made. The partitioning is to be made according to the direction chosen (positive or negative) when each axis direction is sampled for the first time. Thus if \(i+1\) axes are sampled (with \(i + 1 \leq D\)), we have \(\prod_{j=0}^{i} (2D-2j)\) classes of walk each of which produces the same contributions to the mean-square displacement as self-avoiding walks with the following attributes, the first direction sampled is the “1 direction”, that is, \((1, 0, 0, ..., 0)\), etc. The value of \(i+1\) defines the dimensionality of a “sub-lattice” of \(Z^D\) to which the walk is confined. Hence one can write,

\[ C_n \langle \tilde{X}_0^2 \rangle = \sum_{n\text{-stepSAW}} \langle \tilde{X}_0^2 \rangle = \sum_{i=0}^{D-1} R_i \prod_{i=0}^{j} (2D-2j) \]

By the same arguments, one can write

\[ C_n \langle \tilde{X}_0^2 \rangle = \sum_{n\text{-stepSAW}} \langle \tilde{X}_0^2 \rangle = \sum_{i=0}^{D-1} A_i \prod_{i=0}^{j} (2D-2j) \]

To show that \(A_i\) and \(R_i\) are independent of \(D\), if \(i \geq D\) the product term is zero and therefore \(A_i\) and \(R_i\) do not contribute to their respective sums and are therefore unimportant. If \(i < D\), then increasing \(D\) will only change the product term but will leave \(A_i\) and \(R_i\) unchanged since these represent the respective quantities (i.e., sum of the squares of the end-to-end distance and number of walks) after the \(i+1\) dimensions are selected. Though it is obvious that \(R_i\) and \(A_i\) will depend on \(n\) and \(i\).

An expression similar to the ones in the numerator and denominator of Eq. 2 has been stated before without proof for the number of self-avoiding walks. A mathematical proof is important because the method used in the proof of Theorem 2 will help us derive exact expression for \(A_i\) and \(R_i\) in terms of \(n\). In Theorem 4 we derive 9 such values. Also it can be seen from the proof of Theorem 2 that the theorem also holds for simple random walks and also for random walks with exclusion of immediate reversals. In the next theorem we state this and also derive triangular and vertical recursion formulae for the constants therein.

**Theorem 3**

Theorem 2 holds for random walks with exclusion of immediate reversals. Also if \(A_n^{k+1}\) represents \(A_{n+1}^{k+1}\) for \(n\) steps then the following recursion formulae hold for \(A_n^{k+1}\):

\[ A_{n+1,k} = A_{n,k} + (2k-1)A_{n+1,k} \]

\[ A_{n+1,k} = \sum_{j=0}^{n-1} \{ \binom{n-1}{j} A_{n+1,j+1} 2^{n-j-1} \} + A_{n,k} \]

**Proof**

Careful examination of the proof of Theorem 2 reveals that nowhere in the proof have we used the fact that the random walks are self-avoiding. Therefore the theorem holds for random walks with exclusion of immediate reversals and also for simple random walks. To prove (i) we note that \(A_{n+1,k}\) represents the cardinality of the set of random walks of \(n+1\) steps with exclusion of immediate reversals in a ‘k’ dimensional sublattice.
Table 1 - Values of $A_i$ for $n=2$ to $n=14$.

<table>
<thead>
<tr>
<th>n</th>
<th>$A_0$</th>
<th>$A_1$</th>
<th>$A_2$</th>
<th>$A_3$</th>
<th>$A_4$</th>
<th>$A_5$</th>
<th>$A_6$</th>
<th>$A_7$</th>
<th>$A_8$</th>
<th>$A_9$</th>
<th>$A_{10}$</th>
<th>$A_{11}$</th>
<th>$A_{12}$</th>
<th>$A_{13}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>4</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>1</td>
<td>12</td>
<td>9</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>1</td>
<td>35</td>
<td>56</td>
<td>16</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>1</td>
<td>97</td>
<td>304</td>
<td>167</td>
<td>25</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>1</td>
<td>271</td>
<td>1560</td>
<td>1455</td>
<td>391</td>
<td>36</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>1</td>
<td>739</td>
<td>7713</td>
<td>11536</td>
<td>4947</td>
<td>786</td>
<td>49</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>1</td>
<td>2033</td>
<td>37606</td>
<td>86708</td>
<td>55608</td>
<td>13555</td>
<td>1422</td>
<td>64</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>1</td>
<td>5512</td>
<td>180783</td>
<td>629624</td>
<td>581145</td>
<td>203845</td>
<td>31990</td>
<td>2381</td>
<td>81</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>1</td>
<td>15036</td>
<td>866110</td>
<td>4478566</td>
<td>5791591</td>
<td>2807897</td>
<td>618185</td>
<td>67639</td>
<td>3757</td>
<td>100</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>1</td>
<td>40616</td>
<td>4122249</td>
<td>31380281</td>
<td>55861745</td>
<td>36437271</td>
<td>10808922</td>
<td>1630251</td>
<td>131425</td>
<td>5656</td>
<td>121</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>1</td>
<td>110187</td>
<td>19611042</td>
<td>217932262</td>
<td>526755183</td>
<td>453299627</td>
<td>176253722</td>
<td>35189590</td>
<td>3860547</td>
<td>238787</td>
<td>8196</td>
<td>144</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>1</td>
<td>296805</td>
<td>92953924</td>
<td>1502642690</td>
<td>4885668944</td>
<td>5468358236</td>
<td>2732134033</td>
<td>702304343</td>
<td>100679084</td>
<td>8391634</td>
<td>410780</td>
<td>11507</td>
<td>169</td>
<td>1</td>
</tr>
</tbody>
</table>
There are two mutually exclusive ways in which such a walk can be built from a walk of 'n' steps. First by appending a step in a new dimension to a 'n' step walk in a 'k-l' dimensional sublattice. After taking care of symmetry there is only one way in which this can be done and this gives us the $A_{n,k-1}$ term in (i). The second way is to append a step to an 'n' step walk in a 'k' dimensional sublattice in a dimension already sampled before. There are obviously '2k-1' ways of doing this, since we are excluding immediate reversals. This then gives us the $(2k-1)A_{n,k}$ term in (i). Expression (ii) follows from (i). Alternatively both (i) and (ii) can be derived from the fact that the falling factorials $[x]_n$ form a basis set for the vector space of polynomials in 'x' and also the fact that the cardinality of set of 'n' step random walks with exclusion of immediate reversals is equal to $[2D]([2D-1])^{k-1}$ (ref.36).

$A_{n,k}$ can be shown to be equal to$^{37}$:

$$A_{n,k} = 2^{k-1} \sum_{i=0}^{k} (-1)^{k-i} \binom{k}{i} (2i-1)$$

Theorem 4

The following formulae hold for the corresponding $A_i$ and $R_i$, $n \geq i$.

(i) $R_0 = n^2$

(ii) $A_0 = 1$

(iii) $R_1 = n$

(iv) $A_{n,1} = 1$

(v) $R_{n,1} = n(n-1)(n-2) + (n+2)(n-1)$

(vi) $A_{n,2} = (n-1)^2$

(vii) $R_{n,3} = (n-2) \left[ \frac{r(n-3)3n^2-5n+17}{6} + 11 \right]^{-2}$

(viii) $A_{n,3} = (n-3)(n-2)^2(n-3)(n-2)(n-1) + (n-3)(n-2)$

(ix) $A_{n,4} = \frac{(n-4)(n-3)(n-2)^2(n-1)}{6} - \frac{5(n-4)(n-3)(n-2)}{6}$

$\frac{(n-4)(n-3)}{2} - 2(n-4) + 5$

Proof

(i) and (ii) $A_0$ represents the number of walks in exactly one dimension (exactly specified after selection is made) and $R_0$ depends the end-to-end distance squared of this walk and so (i) and (ii) are trivial.

(iii) and (iv) $A_{n,1}$ represents the number of self-avoiding walks of 'n' steps where each step is in a different dimension and $R_{n,1}$ is the end-to-end distance squared of this walk and so (iii) and (iv) are trivial.

(v) and (vi) $A_{n,2}$ represents the number of self-avoiding walks wherein (n-1) steps take any of two directions in n-1 different dimensions and the remaining step can take any of two directions in one of these (n-1) dimensions. So if the remaining step is the last step it can take 2(n-1)-1 on 2n-3 directions, if it is the last but one step it can take 2n-4-1 on 2n-5 directions and so on.

So $A_{n,2} = 2n-3 + 2n-5 + 2n-7 + ... + 1$

$= 2n(n-1)/2 - (n-1)$

$= (n-1)^2$

Alternatively a combinatorial reasoning could be used to arrive at the same result. There are n(n-1)/2 ways of choosing two steps that will have the same dimension. However, the two steps can be either in the same direction or in opposite directions once chosen. So we have a total of [2n(n-1)]/2 ways. But we have to subtract n-1 to take care of immediate reversals and this gives us

$$\frac{2n(n-1)}{2} - (n-1) = (n-1)^2$$

we use a similar combinatorial argument for $R_{n,2}$

$$R_{n,2} = \left[ \frac{n(n-1)}{2} \right], \frac{n(n+2)}{2}$$

$$= \frac{2n(n-1)(n+2)}{2}$$

$$= \frac{n(n-1)(n+2)}{2}$$

(vii) and (viii) the logic used is exactly the same as in
Table 2 - Values of $R_i$ for $n=1$ to $n=14$

<table>
<thead>
<tr>
<th>n</th>
<th>$R_0$</th>
<th>$R_1$</th>
<th>$R_2$</th>
<th>$R_3$</th>
<th>$R_4$</th>
<th>$R_5$</th>
<th>$R_6$</th>
<th>$R_7$</th>
<th>$R_8$</th>
<th>$R_9$</th>
<th>$R_{10}$</th>
<th>$R_{11}$</th>
<th>$R_{12}$</th>
<th>$R_{13}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>4</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>9</td>
<td>16</td>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>16</td>
<td>80</td>
<td>42</td>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>25</td>
<td>327</td>
<td>366</td>
<td>88</td>
<td>5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>36</td>
<td>1208</td>
<td>2590</td>
<td>1182</td>
<td>160</td>
<td>6</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>49</td>
<td>4199</td>
<td>16372</td>
<td>12657</td>
<td>3063</td>
<td>264</td>
<td>7</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>64</td>
<td>14028</td>
<td>96800</td>
<td>119294</td>
<td>45928</td>
<td>6824</td>
<td>400</td>
<td>8</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>81</td>
<td>45533</td>
<td>548088</td>
<td>1038474</td>
<td>597290</td>
<td>136043</td>
<td>13620</td>
<td>592</td>
<td>9</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>100</td>
<td>144612</td>
<td>3014304</td>
<td>8578248</td>
<td>7088988</td>
<td>2322076</td>
<td>347782</td>
<td>25006</td>
<td>828</td>
<td>10</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>121</td>
<td>451300</td>
<td>16234688</td>
<td>68331624</td>
<td>79052027</td>
<td>35783481</td>
<td>7519115</td>
<td>795005</td>
<td>42997</td>
<td>1120</td>
<td>11</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>144</td>
<td>1388484</td>
<td>86104098</td>
<td>530383010</td>
<td>843397868</td>
<td>513470130</td>
<td>145392986</td>
<td>21196754</td>
<td>1664472</td>
<td>70128</td>
<td>1474</td>
<td>12</td>
<td></td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>169</td>
<td>4220575</td>
<td>451236030</td>
<td>4038766876</td>
<td>8711334563</td>
<td>6996234119</td>
<td>2597029232</td>
<td>501337420</td>
<td>53591957</td>
<td>3246383</td>
<td>109514</td>
<td>1896</td>
<td>13</td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>196</td>
<td>12699152</td>
<td>2342479496</td>
<td>30315028332</td>
<td>87820932482</td>
<td>91709527530</td>
<td>43749465898</td>
<td>10876558226</td>
<td>1519784524</td>
<td>124100374</td>
<td>5972478</td>
<td>164910</td>
<td>2392</td>
<td>14</td>
</tr>
</tbody>
</table>
the previous case wherein \( A_{n,3} \) represents the number of self-avoiding walks such that \( n-2 \) steps take any of 2 directions in \( n-2 \) dimensions and the remaining two steps can take any of 2 directions in any of these \( n-2 \) dimensions only. So

\[
A_{n,3} = 1 + \sum_{i=0}^{n-4} \left( i+1 \right) + \sum_{j=0}^{33} \left( i+1 \right) + \sum_{k=1}^{120} \left( i+1 \right)
\]

\[
= 1 + 11(n-3) + 10(n-3)(n-4) + 11(n-3)(n-4)(2n-7)/6 + (n-4)^2/2
\]

\[
= \frac{(n-3)(n-2)(n-1)(3n-2)}{6} + \frac{(n-3)(n-2)}{2} + 1
\]

Alternatively a combinatorial argument outlined above can be used to arrive directly at

\[
\frac{n(n-1)(n-2)(n-3)- (n-1)(n-2)(n-3)- n(n-1)(n-3)}{4n(n-1)(n-2)}
\]

\[
\frac{(n-3)(n-2)(n-1)(3n-2)}{6} + \frac{(n-3)(n-2)(n-1)}{2} + 1
\]

In the step (a) the first term represents the selection of 2 pairs with exclusion of immediate reversals, the second and third terms represent selection of a triplet with exclusion of immediate reversals and the last term represents exclusion of 4 step reversals.

A combinatorial argument similar to the one used in (viii) and above can be used to show

\[
R_{n,3} = \frac{n(n-3)(3n^2-5n+17)}{6} + 11
\]

The intermediate step in this case is

\[
R_{n,4} = \frac{n(n-1)(n-2)(n-3)(n+4)}{8}
\]

Theorem 4

The number of \( 2n \) step self-avoiding polygons \( q_{2n} \) in \( D \) dimensional Euclidean space is given by the following formula:

\[
q_{2n} = (2D)^2 \prod_{i=1}^{n-1} \left( \frac{(2D)-2j}{2} \right)
\]

where \( P_i \) is independent of \( D \).

Proof

For a self-avoiding polygon of \( 2n \) steps if any \( 'n' \) steps take any of \( 'n' \) predetermined directions, then the remaining \( 'n' \) steps would be in a direction exactly opposite to the first \( 'n' \) steps (without violating the self-
avoidance criterion of course). The logic of Theorem 2 would therefore be applicable in this case also.

**A New Problem in Graph Theory**

Finally we use the self-avoiding walk problem to define a new problem in Graph theory. The reader is referred to commonly available texts by Harary, Deo and König for the standard terminology of Graph theory.

Consider a rooted tree whose edges are chosen from any of '2D' colours. Also assume that the colours are present in pairs, i.e., each colour has a complement, for example black and white, etc. At each vertex of the tree any of the '2D' colours can be chosen provided at all vertices the following condition is met. In the path (there is only one such path) from the root to the vertex in question it should **not** be possible to identify a subpath in which it is possible to assign to each edge in the subpath an edge of the complementary colour (a subpath can be defined as a proper subset of a path containing adjacent edges only). Then in such a tree "how many pendant vertices of distance 'n' from the root are present?" is the question to be answered. Note: The radius of the tree is not necessarily 'n' due to the presence of traps such as 'rruulldr' in an 8 step two dimensional walk ('r' stands for right, 'l' for left, 'u' for up and 'd' for down).

Efforts are on to solve the graph theoretic version of this problem. In the meanwhile the definition of the graph theoretic problem has proven to be extremely useful in designing some computer algorithms that can calculate the coefficients $A$, $R$ and $P$ in Theorems 2, 4 and 5. Computer algorithms for $A$, $R$ and $P$ are discussed below. The reader is directed here to Manber, Hare and Knuth for the Analysis of Algorithms and Hennessey for details on Common LISP implementation.

**Algorithm I - Divide and conquer**

A tree of 2n steps is built up by appending at each pendant vertex of an 'n' step tree a copy of the same tree and the condition mentioned in the problem is checked for cases that it has not been checked before (the two copies of the 'n' step tree already obey the condition). This process is repeated recursively. For a (2n+1) step tree a copy of the 'n' step tree is taken a single edge appended to each pendant vertex, violation of the condition checked for, and the new copy of the (n+1) step tree appended to the original 'n' step tree. This is then repeated recursively.

To calculate the coefficients $A$, $R$ and $P$ one can calculate the relevant quantities (number of self-avoiding walks for $A$, sum of the squares of the end-to-end distance for $R$, etc.) for all dimensionalities up to 'n' the number of steps and then invert the resulting simultaneous equations using Cramer’s rule. This is, however, not necessary as the symmetry in Theorems 2, 4 and 5 could be incorporated into the Divide and Conquer algorithm, using a standard binary search algorithm, and the time required considerably shortened. This uses a Divide and Conquer technique to enumerate the coefficients in Theorems 2, 3 and 4 without exploiting the symmetry fully and finally using Crammer’s rule. The final program in 'C' that calculates the coefficients in Theorems 2, 4 and 5 using the Divide and Conquer technique directly is a lot more complicated and confusing and hence is not being discussed here. It will be provided to the interested reader on request and on receipt of a floppy or an e-mail address.

<table>
<thead>
<tr>
<th>$2n$</th>
<th>$P_1$</th>
<th>$P_2$</th>
<th>$P_3$</th>
<th>$P_4$</th>
<th>$P_5$</th>
<th>$P_6$</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>3</td>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>14</td>
<td>62</td>
<td>27</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>70</td>
<td>970</td>
<td>1205</td>
<td>248</td>
<td></td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>372</td>
<td>15691</td>
<td>44885</td>
<td>24105</td>
<td>2830</td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>2058</td>
<td>263011</td>
<td>1612394</td>
<td>1750966</td>
<td>507892</td>
<td>38232</td>
</tr>
</tbody>
</table>
Table 4 - The values of the ratio $R_i/A_i$ for $n = 2$ to 14

<table>
<thead>
<tr>
<th>$n$</th>
<th>$R_i/A_1$</th>
<th>$R_i/A_2$</th>
<th>$R_i/A_3$</th>
<th>$R_i/A_4$</th>
<th>$R_i/A_5$</th>
<th>$R_i/A_6$</th>
<th>$R_i/A_7$</th>
<th>$R_i/A_8$</th>
<th>$R_i/A_9$</th>
<th>$R_i/A_{10}$</th>
<th>$R_i/A_{11}$</th>
<th>$R_i/A_{12}$</th>
<th>$R_i/A_{13}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>4</td>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>6.670</td>
<td>4.667</td>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>9.343</td>
<td>6.536</td>
<td>5.500</td>
<td>5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>12.454</td>
<td>8.520</td>
<td>7.078</td>
<td>6.400</td>
<td>6</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>15.494</td>
<td>10.495</td>
<td>8.699</td>
<td>7.834</td>
<td>7.333</td>
<td>7</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Algorithm II - Back-tracking

Standard back-tracking\textsuperscript{23,24} technique can be applied to the graph theoretic problem mentioned above. Here at each vertex a single edge is checked for violation of the condition. If the condition is violated a new edge is chosen, if not the process continues until a pendant vertex (A self-avoiding walk) is reached. The process is then repeated recursively by back-tracking one edge at a time. To evaluate the coefficients $A_n$, $R_n$ and $P_n$, one could calculate the relevant quantities for all dimensionalities up to $n$ the number of steps and then invert the resulting simultaneous equations using Cramer's rule. However, as mentioned in the case of 'divide and conquer' it would be much easier to exploit the symmetry of Theorems 2, 4 and 5 and include this in the back-tracking algorithm itself. Sample programs in C and in COMMON LISP along with a detailed description will be provided on request and on receipt of a floppy or an e-mail address.

Algorithm III - The first step append algorithm

At the root of an 'n' step tree an edge of a predetermined colour (due to symmetry we do not need to consider the other 2D-1 colours) is appended recursively. We then compare the first $n$ steps of the new tree (i.e., the new tree with $(n+1)^{th}$ edges excluded) with the old tree. Any branches in the new tree not present in the old tree are then removed. If the number of steps is even a return to the origin in all cases is checked for, if the number is odd no checking is required. This then gives us the number of self-avoiding walks because steps 2 to $(n+1)$ are self-avoiding by construction whereas steps 1 to $n$ are self-avoiding by the comparison process described above. So only a return to the origin has to be checked for in even cases.

This algorithm is in theory much faster as it involves only one checking in even cases and no checking in odd cases. However, it suffers from a major draw-back that it is exponential in memory requirement and hence difficult to exploit in its current form. A sample program in C will be provided on request and on receipt of a floppy or an e-mail address.

Evaluation of $A_n$ (for 'n' up to 14), $R_n$ (for 'n' up to 14) and $P_n$ (for '2n' up to 14)

We have evaluated $A_n$, $R_n$ (for 'n' up to 14) and $P_n$, (for 'n' up to 14). The values are given in Tables 1, 2 and 3. The values of $A_n$ for $n = 2$ to 11 have been reported before\textsuperscript{26,27} and are included here for completeness only.

Comparison with exact enumeration data

To check if the formulae in Theorems 2 to 4 agree with computer enumeration data one has to check against values of all $D$ up to $D = n$ (i.e., $D = n$ in case self-avoiding polygons of size $2n$). Many of the values are readily available in the literature\textsuperscript{22,25} and are available in consolidated form in Ref. 25. When not available the authors have carried out the exact enumeration themselves in all the cases required. Except for three obvious typographical errors in the data in Ref. 25 our formulae are found to agree perfectly in all cases.

Conclusion

In this paper we have provided an exact formula for $(\langle n(n)\rangle^2)$ for simple random walks with exclusion of immediate reversals and also exact formulae for $(\langle n^2(n)\rangle^2)$ for self-avoiding walks up to $n=14$. We have also derived some of the constants involved in these formulae for all $n$. We have also described a new problem in Graph Theory and used the description of this problem to design three different types of computer algorithms.

Acknowledgement

The authors would like to thank Prof. K N R Nair of the Computer Science Department of Mahatma Gandhi University for many helpful discussions and also for providing the fundamental program that computed the number of self-avoiding walks, a modification of which formed basis of the back-tracking type programs that computed the constants for Theorem 2. A S P would also like to thank the Mathematical Sciences Division, DST, New Delhi for funding this work.

References