

RANDOM WALK STUDIES IN TWO DIMENSIONS*

G. Luste and R. Blankenbecler
Stanford Linear Accelerator Center
Stanford University, Stanford, California 94305

ABSTRACT

Computer studies of random walks in two dimensions are reported. The concept of degeneracy D is introduced in which only D steps are allowed at any one lattice point. Properties of the probability distribution function for the distance covered in N steps are discussed. The lifetime of a walk of degeneracy D is also examined.

(Submitted to Journal of Mathematical Physics.)

*Work supported by the U. S. Atomic Energy Commission.

I. Introduction

The study of the statistical properties of random walks has had a considerable history¹ because of its intrinsic mathematical interest and its application to a variety of physical problems. This note will be restricted to a study of such walks on a two dimensional square grid. The classical (and most familiar) unrestricted random walk permits any number of steps on any lattice point. In the limit of a large number of steps, this problem gives rise to the familiar Gaussian distribution of end points.² Certain physical applications involve a random walk problem but with restrictions on the steps allowed.^{3,4} In the simple theories of the size of long polymer molecules, the problem of the excluded volume is taken into account by forbidding more than one step at any lattice point.⁵ This type of restriction also occurs in theories of particle production in high energy collisions which was the original motivation behind the present work.⁶

Because such walks do not at present have a complete analytic solution, we have relied extensively on computer simulations in this study. Hopefully, the results presented here may be of use in future physical applications and may also suggest directions to be explored analytically.

Our main concern will be with random walks on a square unit lattice and the distribution function for the distance achieved from the origin after N steps. The study of walks allowing a maximum of $2, 3, 4, \dots$ steps on any one lattice point represents a natural extension of the excluded volume walk. We will therefore introduce the concept of degeneracy denoted by the integer D . By definition, in a random walk of degeneracy D , it is forbidden to step more than D times on any one lattice point. The classic free random walk corresponds to an infinite degeneracy, and the excluded volume problem has $D = 1$.

It will be shown that the random walks for small D ($D \leq 4$) are remarkably similar and have a simple dependence on D . However, the finite degeneracy case is very different from the unrestricted or infinite D case. It must be kept in mind that our numerical, or better, experimental, analysis is limited to $N \lesssim 50$ steps. Thus any statement that requires an extrapolation to infinite N is dependent upon an assumption that we are dealing with a rapidly convergent variable.

In the next section, the numerical procedures are discussed and an outline of the remainder of the paper is presented.

II. Numerical Procedure and Analysis

The numerical results were determined by examining a large sample of random walks simulated on the SLAC IBM 360-91 computer.⁷ Each step in a walk is random to the extent that a series of pseudorandom numbers can be generated. In practice, this is not a problem. A generalized random walk is easily constructed by introducing a two dimensional array $M(ij)$ and setting all locations equal to zero. The walk starts at the center location or lattice point. Whenever a "step" is taken to the point ij , the number one is added to $M(ij)$ for that particular location. Thus $M(ij)$ records the number of steps taken at each lattice point. The first time an element of $M(ij)$ exceeds D , the walk is terminated and the process is repeated for another walk.⁸

In taking a step, the computer is asked to randomly select one of the four allowed directions on this grid. The resulting step is then taken and $M(ij)$ is incremented as is the step count N . Figure 1 shows an example of such a walk for $D = 1$ and $N = 15$ steps. For each walk, the values of r_N , r_N^2 , $r_N(\text{MAX})$, etc. are recorded. Statistical distributions in these variables can be generated by performing

a large number of uncorrelated random walks. Functional fits as well as possible correlations can then be studied. The main quantities of interest which are to be studied as functions of D and N , the total number of steps in a walk, are:

- r_N — the distance from the origin to the endpoint at step N .
- $\langle r_N \rangle$ — the mean r_N value for a large number of walks of N steps.
- T_N — the fractional number of walks that are terminated on step $N + 1$ by the degeneracy constraint.

Two other quantities which will be briefly discussed are

- \bar{r}_N — the average distance from the origin of each step in a walk of length N .
- $r_N^{(\text{MAX})}$ — the maximum r_i ($i=1, \dots, N$) in a walk of N steps.

In Section III the dependence of $\langle r_N^2 \rangle$ on N and on D is examined. Comparisons with Fisher ($D=1$) and with the known dependence of $\langle r_N^2 \rangle = N$ for the unrestricted walk ($D = \infty$) are made.

In Section IV the functional form of the r_N distribution is studied. It will be shown to approximate a simple displaced Gaussian, which scales as $\langle r_N^{(D)} \rangle$.

Section V looks at $T_N^{(D)}$ and for large N fits it with an exponential. The D dependence is also shown. A few additional comments are made in Section VI.

This study emphasizes the gross behavior of the above variables and does not claim to understand all the fine structure. The tables list the "experimental" numbers in detail. The figures however emphasize the qualitative features of the data.

III. The Mean Square Displacement

Let us consider properties of the mean square displacement from the origin after N steps, defined by $\langle r_N^2(D) \rangle$. Following Wall and collaborators,⁷ and Hiley and Fisher,⁵ it will be assumed that $\langle r_N^2(1) \rangle$ varies as a power of the number of steps N if N is large. To explore the D dependence of this quantity, we compute the ratio

$$R_N(D) = \langle r_N^2(1) \rangle / \langle r_N^2(D) \rangle . \quad (1)$$

If $R_N(D)$ becomes constant for large N , then the power behavior of the mean square radius is independent of D . The dependence of $R_N(D)$ can then be easily extracted. One might expect a simple dependence of D from the following qualitative argument. A value of $D \geq 2$ can be considered to be a $D = 1$ random walk on D parallel planes. Thus one could expect that the value of $\langle r_N^2 \rangle$ achieved after many steps to be smaller by a factor of D . (This will not be the case for large D at fixed N since the resultant mean square radius cannot be smaller than N , which is the value expected for infinite degeneracy.)

Using the computer procedure described above, the values of $\langle r_N^M(D) \rangle$ were computed for $D = 1, 2, 3, 4$, and ∞ and $M = 1, 2, 3$, and 4 . Numerical values are presented in Tables I and II. In Figure 2, the ratio $R_N(D)$ of Eq. (1) is given as a function of N . The results are very suggestive that $R_\infty(D) \cong D$ and also that the exponent of the power behavior of $\langle r_N^2(D) \rangle$ as a function of N is independent of D (for small D). Using this result, it was assumed that

$$\langle r_N^2(D) \rangle = \frac{A}{D} N^\gamma + \dots , \quad (2)$$

and one finds that $A \simeq 0.80$ and $\gamma \simeq 1.5$. The values of γ that were extracted from the data for $D = 1$ using neighboring N points are shown on Figure 3. In Figure 3 the dotted line uses the extrapolation procedure of Ref. 5, that is the first term of Eq. 2, to extract γ . This procedure is extremely sensitive and yields large fluctuations in γ for $N \geq 10$. The solid line in Figure 3 is a different extrapolation procedure which is an attempt to smooth out these fluctuations. These results are in good agreement with the values given by Refs. (5) and (7) who fit Eq. (2) for the case $D = 1$ and found $\gamma = 1.47 \pm .02$ for $N = \infty$.

A more specific assumption concerning the behavior of the mean square radius which could be tested is that

$$\langle r_{N(D)}^2 \rangle = NG(N/D^2).$$

The value $G(0) = 1$ follows from the known value in the case of infinite degeneracy. One also expects the G is larger than one for finite argument and to increase like a power $(\gamma - 1)$ for large argument. This form is consistent with all our results but the tests are not conclusive for $D = 3$ and 4 .

IV. Scaling of the Distribution Function

Now we shall define tests to see if the end-point distribution function scales. Scaling for a fixed D is defined in the following way. If the probability $P(r, N)$ of ending up in rdr after N steps is a function of $r/F(N)$ and there is no other N dependence (except for the normalization constant which must go as $F^{-2}(N)$), then the distribution will be said to scale. This property in turn implies a simple behavior for ratios of moments of the distribution. It is convenient to introduce in the scaling case

$$P(r, N) = F^{-2}(N) Q(r/F(N)) \quad (3)$$

and then

$$1 = \int_0^{\infty} r dr P(r, N) = \int_0^{\infty} y dy Q(y) .$$

Hence it follows from scaling that the ratios

$$\langle r_N \rangle / \langle r_N^M \rangle^{1/M} \quad (4)$$

are independent of N and depend only on M . Evidence will be given that the distribution functions scale for the generalized random walks considered here.

Numerical results for this ratio as a function of N and D for several M values are listed in Tables I and II. It is seen that for N greater than fifteen or so, the ratios seem to be approaching constant values. We then infer that the distribution functions scale. Furthermore, the values of these ratios for $D = 1, 2,$ and 4 imply that their distribution function is pushed further away from the origin than the known $D = \infty$ Gaussian case. This is as one would expect.

To pursue this point further, the distribution function for $D = 1$ and the sum of $N = 23$ and 24 is shown in Figure 4. These two values of N were combined in order to smooth out the fluctuations expected from even and odd values of N and the effects of the discrete grid. Such fluctuations are evident in Figure 3 and are not due just to statistical effects. Several different analytical forms have been used to fit this distribution. A good fit is achieved with the form

$$r dr P(r, N) = y dy Q_1 \exp [-a(y-b)^2] , \quad (5)$$

where

$$y = r / \langle r_N \rangle ,$$

and the fitted parameters are $a = 2.8$ and $b = 0.79$. The parameters a and b are not independent but are related by the fact that $\langle r_N \rangle$ provides an identity. The fit for this distribution is shown as the solid line in Figure 4. This form also fits quite well down to $N \sim 12$ and steadily improves as N increases.

A function that provides a slightly better fit at small N is

$$P(r, N)rdr = dy Q_0 \exp(-c(y-d)^2), \quad (6)$$

where $c = 3.6$ and $b = 1.0$. This probability function is not proportional to the area element (rdr). These two fits are compared in Figure 4.

In Figure 5, scaling is illustrated for $D = 1$ and selected N values between 11 and 41. The best χ^2 fit using Eq. (5) is shown as the solid line in this figure and the respective values of a and b are given in Table III. The constancy of the values of a and b is further proof of scaling for $D = 1$.

The function given by Eq. (5) was also fitted to the experimental distribution functions for $D = 2, 3, 4$ and ∞ . The values found for a, b , and the corresponding $\langle r \rangle$ are given in Table III. One knows that for $D = \infty$, $a = \pi/4 = 0.785$, and $b = 0$. This is quite consistent with our results. It should be noted that the values of $r < 2$ in the data contribute a large χ^2 to this fit, especially if $\langle r \rangle$ is small. It is just this region where one expects the finite structure of the grid to be significant. If $r < 2$ is omitted in the fit to the data, the values of a and b increase, in some cases as much as ten percent. For example, for $D = \infty$, a is increased to 0.78 and b to + 0.02.

The values of a and b are very constant for $D = 1, 2$, and ∞ . For $D = 3$ and 4, we conjecture that the monotonic increase in b is due to our limited N values. As D increases, the asymptotic limit for b is reached only at larger N values. However, for $D = \infty$, the initial value, as well as the limit value, is equal to zero.

V. Step Lifetime

Let $C_N(D)$ be the total number of distinct random walks of N steps available on a two dimensional grid of degeneracy D . The probability that a walk starting at the origin will survive to the N^{th} step is

$$P_N(D) = C_N(D)/C_N(\infty) ,$$

where $C_N(\infty) = 4^N$. The probability that a walk again starting at the origin will be terminated (by the D constraint) at the $(N+1)^{\text{st}}$ step is

$$T_N(D) = P_N(D) - P_{N+1}(D) .$$

The variation of $T_N(D)$ for $D = 1, 2$, and 3 is shown on a semilog scale in Figure 6.

At larger N , an obvious exponential dependence is observed. An estimated straight edge fit to the form $T_N(D) = \exp(-N/\lambda(D))$ gives $\lambda(1) \cong 2.6$, $\lambda(2) \cong 8.3$, and $\lambda(3) \cong 20.0$. The insert in Figure 6 suggests that $\lambda(D)$ may be equal to $\exp(D)$. In physical terms, $\lambda(D)$ is the average lifetime (in terms of the number of steps) for a walk of degeneracy D . For small N , T_N does not have this simple exponential behavior since it must vanish for N less than $2D$.

Fisher and Sykes⁴ have presented evidence that the total number of non-self-interesting random walks has the form

$$C_N(1) \approx B N^\alpha \mu^N .$$

For some very interesting rigorous results on $C_N(1)$ see references 9 and 10. Our fit yields $\mu \sim 2.4$ which is in reasonable agreement with the value 2.64 given by Hiley and Fisher.⁵

VI. Additional Comments

There are several additional interesting questions that can be asked about random walks which have physical application. The quantity $\langle r_N \rangle$ measures the average displacement achieved between the first and the last step. Let us also consider the quantity $\langle \bar{r}_N \rangle$, which is defined as the average distance from the origin to all covered lattice points. It is the average of the distances r_i illustrated in Figure 1. This quantity is given in Table IV for $D = 1$ and ∞ .

It is also of some interest to know the maximum distance achieved from the origin during the course of a walk. This quantity, $\langle R_N(\text{MAX}) \rangle$, is illustrated in Figure 1 and is also given in Table IV. It is seen to vary with N in the same way as $\langle r_N \rangle$. For $D = 1$, $\langle R_N(\text{MAX}) \rangle$ is ten percent larger than $\langle r_N \rangle$ but for $D = \infty$, it is twenty-five percent larger.

Footnotes and References

1. For an excellent review see M. E. Fisher, Reports on Progress in Physics XXX, Part II, 1967.
2. S. Chandrasekhar, Rev. Mod. Phys. 15, 3 (1943).
3. E. W. Montroll, J. Chem. Phys. 18, 734 (1950).
4. M. E. Fisher and M. F. Sykes, Phys. Rev. 114, 45 (1959).
5. B. J. Hiley and M. E. Fisher, J. Chem. Phys. 34, 1253 (1961).
6. L. Stodolsky, SLAC-PUB-864. Application of these results to a generalized multiperipheral model will be discussed in the near future by T. Neff, R. Savit and R. Blankenbecler (to be published).
7. See also the article "Monte Carlo Methods Applied to Configurations of Flexible Polymer Molecules" by F. T. Wall, S. Windwer, P. J. Cans, Methods in Computational Physics 1, 217 (1963).
8. A typical computer run takes two minutes for a study of 10^4 different walks and we were thus limited to a sample of $\leq 10^5$ walks.
9. J. M. Hammersley, Proc. Camb. Phil. Soc. 53, 642 (1957).
10. H. Kesten, J. Math. Phys. 5, 1128 (1964).

TABLE I

Excluded Volume Random Walks (D=1)

N	$\langle r \rangle$	$\pm \Delta r$	$\langle r^2 \rangle$	$\pm \Delta r^2$	$\frac{\langle r \rangle}{\sqrt{\langle r^2 \rangle}}$	$\frac{\langle r \rangle}{\sqrt[3]{\langle r^3 \rangle}}$	$\frac{\langle r \rangle}{\sqrt[4]{\langle r^4 \rangle}}$
1	1.000	-	1.00	-	1.000	1.000	1.000
2	1.611	0.001	2.67	-	0.986	0.971	0.957
3	2.047	0.002	4.56	0.01	0.959	0.929	0.907
4	2.557	0.002	7.04	0.01	0.964	0.934	0.911
5	2.951	0.003	9.56	0.02	0.954	0.921	0.894
6	3.392	0.004	12.58	0.02	0.956	0.923	0.896
7	3.748	0.004	15.55	0.03	0.950	0.914	0.885
8	4.150	0.005	19.02	0.04	0.952	0.915	0.886
9	4.480	0.006	22.35	0.05	0.948	0.909	0.879
10	4.849	0.007	26.14	0.07	0.949	0.910	0.879
11	5.169	0.008	29.86	0.08	0.946	0.906	0.875
12	5.516	0.009	34.0	0.1	0.946	0.907	0.875
13	5.83	0.01	38.1	0.1	0.944	0.904	0.872
14	6.16	0.01	42.5	0.1	0.945	0.904	0.872
15	6.46	0.01	46.9	0.2	0.944	0.902	0.870
16	6.79	0.01	51.7	0.2	0.944	0.903	0.870
17	7.07	0.02	56.2	0.2	0.943	0.901	0.867
18	7.36	0.02	61.0	0.3	0.943	0.900	0.866
19	7.65	0.02	66.0	0.3	0.942	0.899	0.865
20	7.95	0.02	71.2	0.3	0.942	0.899	0.865
21	8.21	0.02	76.2	0.4	0.941	0.897	0.862
22	8.50	0.03	81.6	0.4	0.941	0.897	0.862
23	8.76	0.03	86.8	0.5	0.940	0.896	0.861
24	9.04	0.03	92.3	0.5	0.941	0.897	0.862
25	9.29	0.03	97.7	0.6	0.940	0.896	0.861
30	10.74	0.05	131.0	1.0	0.939	0.895	0.861
35	11.99	0.08	164.0	2.0	0.935	0.888	0.852
40	13.3	0.1	201.0	3.0	0.939	0.894	0.858
45	14.6	0.2	242.0	5.0	0.940	0.895	0.859
50	15.8	0.3	283.0	8.0	0.938	0.892	0.856

TABLE II

Random Walks with $D = 2, 4, \infty$

N	D = 2			D = 4			D = ∞		
	$\langle r \rangle$	$\langle r^2 \rangle$	$\frac{\langle r \rangle}{\sqrt{\langle r^2 \rangle}}$	$\langle r \rangle$	$\langle r^2 \rangle$	$\frac{\langle r \rangle}{\sqrt{\langle r^2 \rangle}}$	$\langle r \rangle$	$\langle r^2 \rangle$	$\frac{\langle r \rangle}{\sqrt{\langle r^2 \rangle}}$
1	1.000	1.00	1.000	1.000	1.00	1.000	1.000	1.00	1.000
2	1.308	2.00	0.854	1.207	2.00	0.853	1.207	2.00	0.853
3	1.587	2.99	0.917	1.587	3.00	0.917	1.588	3.00	0.917
4	1.869	4.26	0.905	1.753	4.00	0.877	1.752	4.00	0.876
5	2.160	5.54	0.917	2.020	5.00	0.903	2.024	5.02	0.903
6	2.423	7.01	0.915	2.156	5.98	0.881	2.164	6.02	0.882
7	2.676	8.49	0.919	2.372	6.99	0.897	2.379	7.02	0.898
8	2.930	10.16	0.919	2.508	8.01	0.886	2.503	8.00	0.885
9	3.159	11.79	0.920	2.691	9.02	0.896	2.685	9.00	0.895
10	3.400	13.62	0.920	2.826	10.10	0.889	2.800	9.98	0.886
11	3.62	15.44	0.921	2.995	11.18	0.896	2.962	10.99	0.893
12	3.86	17.49	0.922	3.127	12.30	0.892	3.067	11.97	0.886
13	4.07	19.47	0.922	3.284	13.41	0.897	3.215	12.99	0.892
14	4.29	21.60	0.923	3.408	14.55	0.893	3.316	13.99	0.887
15	4.50	23.8	0.924	3.550	15.67	0.897	3.448	14.98	0.891
16	4.72	26.1	0.925	3.671	16.85	0.894	3.539	15.95	0.886
17	4.92	28.3	0.924	3.817	18.07	0.898	3.661	16.93	0.890
18	5.12	30.7	0.925	3.935	19.30	0.896	3.751	17.92	0.886
19	5.31	33.0	0.926	4.072	20.55	0.898	3.870	18.92	0.890
20	5.50	35.3	0.926	4.188	21.82	0.897	3.956	19.90	0.887
21	5.70	37.8	0.927	4.317	23.07	0.899	4.068	20.90	0.890
22	5.89	40.3	0.927	4.431	24.36	0.898	4.150	21.90	0.887
23	6.08	43.0	0.928	4.557	25.65	0.900	4.253	22.87	0.889
24	6.28	45.8	0.928	4.671	26.97	0.899	4.333	23.87	0.887
25	6.46	48.5	0.928	4.791	28.27	0.901	4.435	24.91	0.889
30	7.32	62.1	0.930	5.37	35.3	0.903			
35	8.19	78.0	0.928	5.93	42.9	0.905			
40	9.3	101.0	0.928	6.49	51.5	0.905			
45	10.5	130.0	0.923	7.02	60.2	0.904			
50	11.6	158.0	0.924	7.54	69.5	0.905			

TABLE III
 Distribution Function $y e^{-a(y-b)^2}$ Fits

Degeneracy	N Steps	a	b	$\langle r \rangle$	# Bins
D = 1	5+6	2.5	.76	3.13	6
	11+12	2.7	.79	5.31	12
	17+18	2.8	.80	7.18	17
	23+24	2.8	.79	8.87	20
	31+32	2.5	.76	11.10	23
	40+41	2.7 \pm .1	.78 \pm .01	13.50	28
D = 2	23+24	2.0	.67	6.12	16
	31+32	2.1	.70	7.41	18
	40+41	2.1 \pm .1	.69 \pm .02	9.44	20
D = 3	23+24	1.2	.42	5.05	14
	31+32	1.4	.51	6.15	17
	40+41	1.5	.55	7.41	21
D = 4	23+24	.92	.17	4.47	14
	31+32	1.09	.33	5.51	16
	40+41	1.00 \pm .01	.40	6.09	20
D = ∞	23+24	.75	-.05	4.26	14
	31+32	.74	-.04	4.95	16
	40+41	.74	-.06	5.61	18

TABLE IV

Random Walks with $D = 1, \infty$

D = 1					
N	$\langle r \rangle$	$\langle \bar{r} \rangle$	$\langle r_{MAX} \rangle$	$\langle r^2 \rangle$	$\langle \bar{r}^2 \rangle$
9	4.48	2.94	4.84	22.4	10.9
10	4.85	3.15	5.24	26.1	12.6
15	6.46	4.13	7.09	46.9	22.0
16	6.79	4.31	7.45	51.7	24.1
22	8.50	5.39	9.47	81.6	37.9
23	8.76	5.55	9.78	86.8	40.3

D = ∞					
9	2.69	1.90	3.18	9.00	4.98
10	2.80	1.99	3.37	10.0	5.48
15	3.45	2.39	4.20	15.0	7.99
16	3.54	2.46	4.35	16.0	8.52
22	4.15	2.85	5.16	21.9	11.5
23	4.25	2.91	5.28	22.9	12.0

FIGURE CAPTIONS

1. Definition of the quantities of interest for a specific walk.
2. The ratio test for the D dependence of the mean square radius.
3. The power behavior of the mean square radius for $D = 1$.
4. The probability distribution function for $N = 23 + 24$ steps and the fits as described in the text.
5. The scaling properties of the probability distribution function for $D = 1$.
6. The number of walks terminated immediately after the N^{th} step by the D constraint.

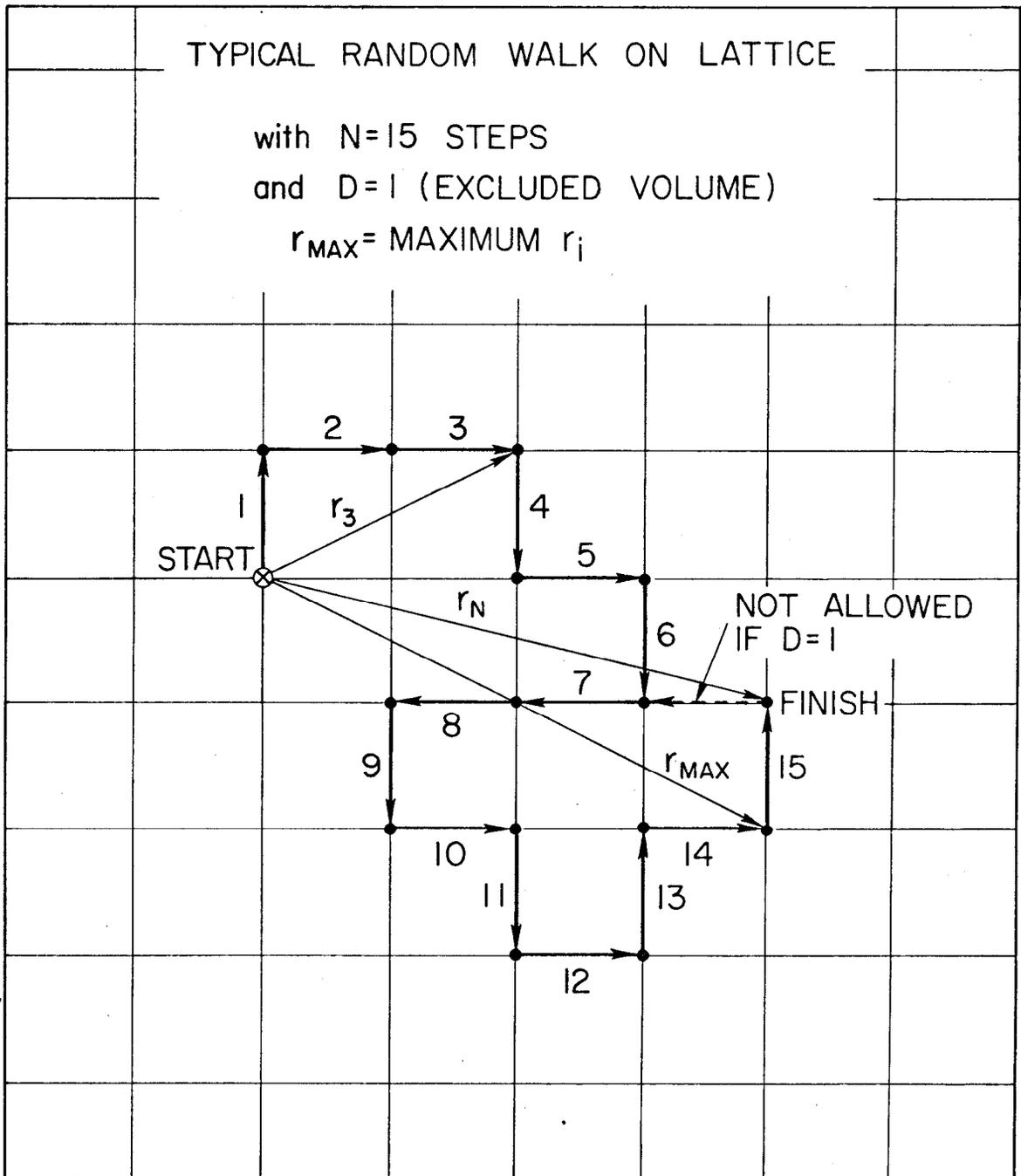


Fig. 1

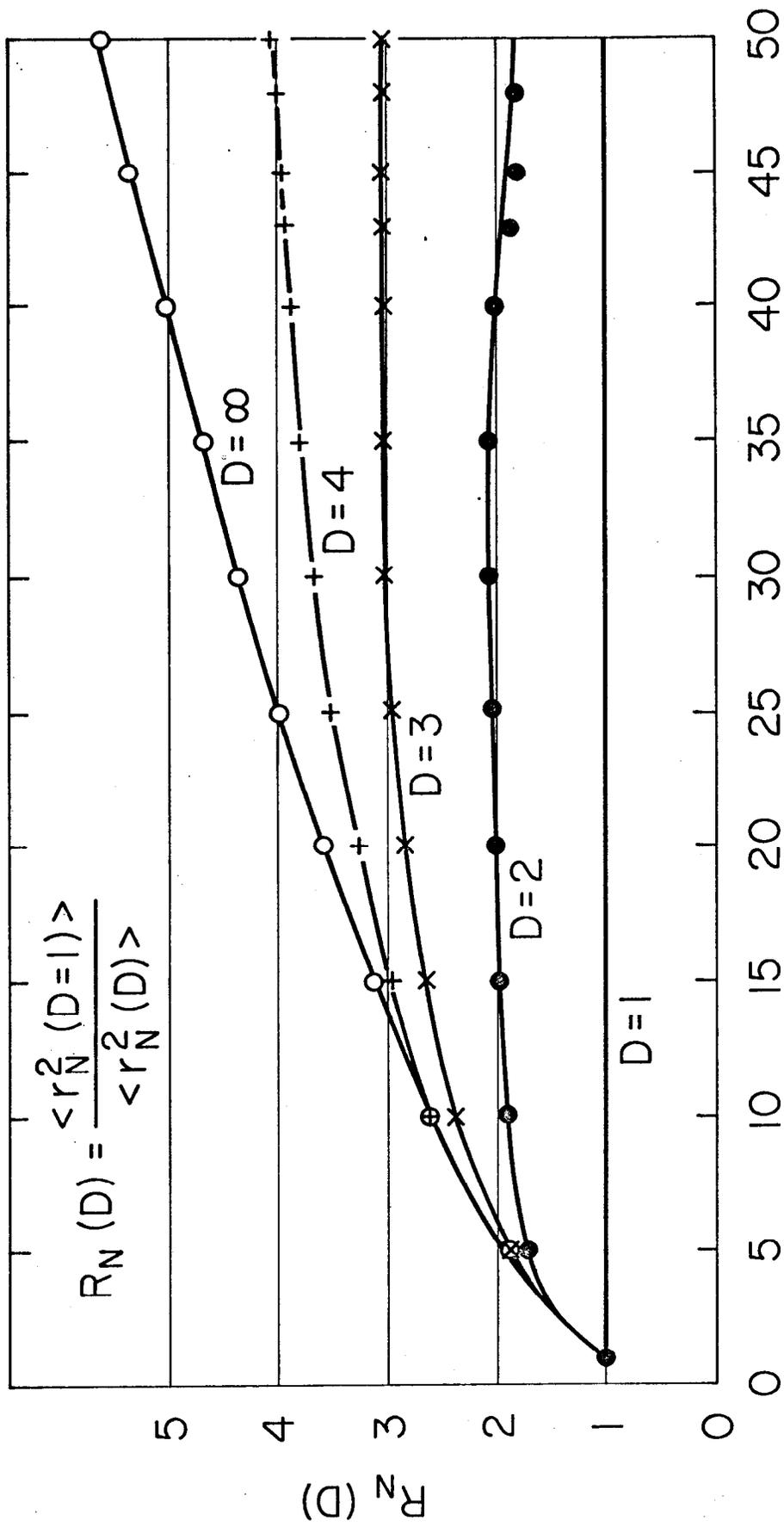


Fig. 2

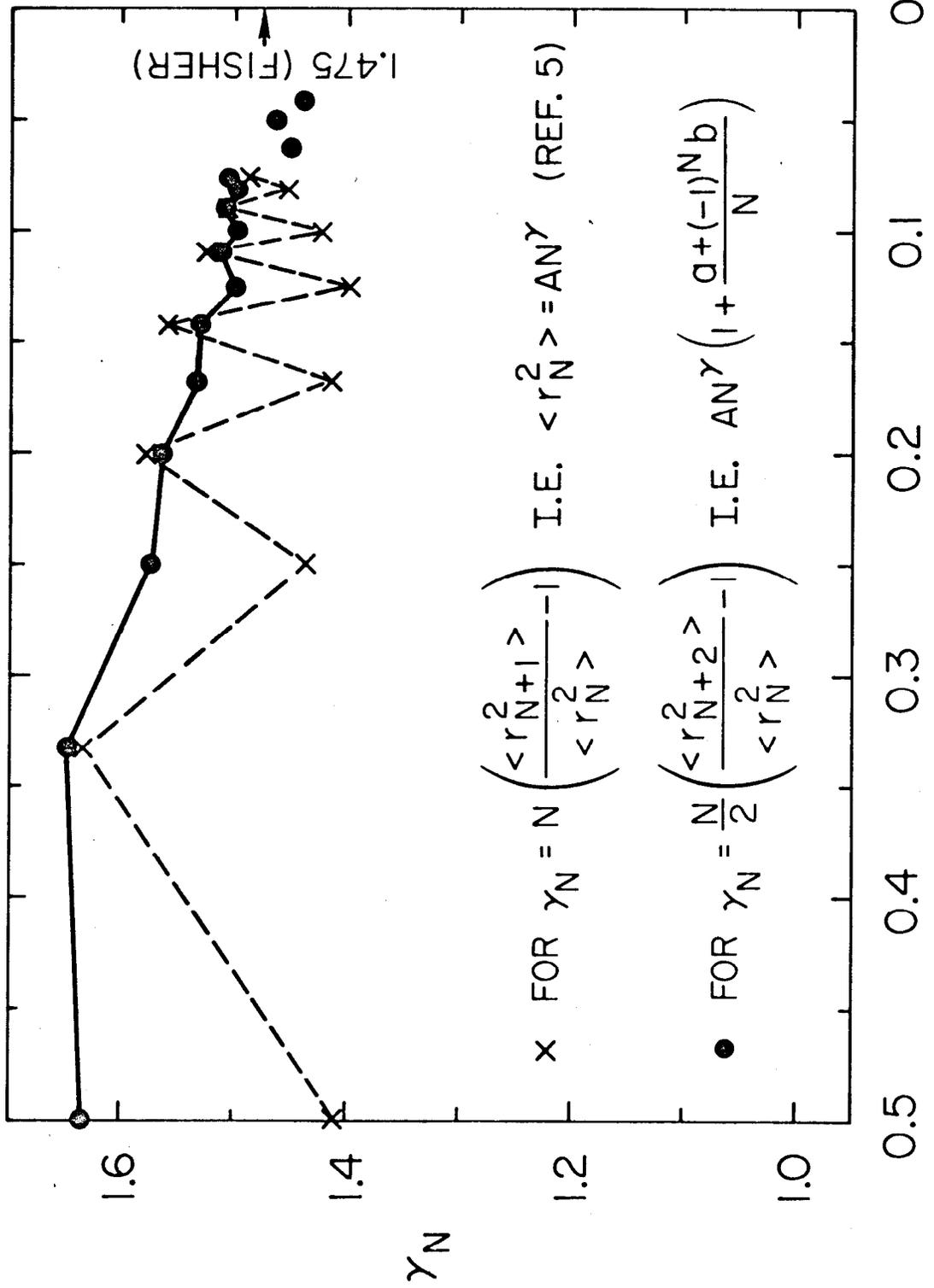


Fig. 3

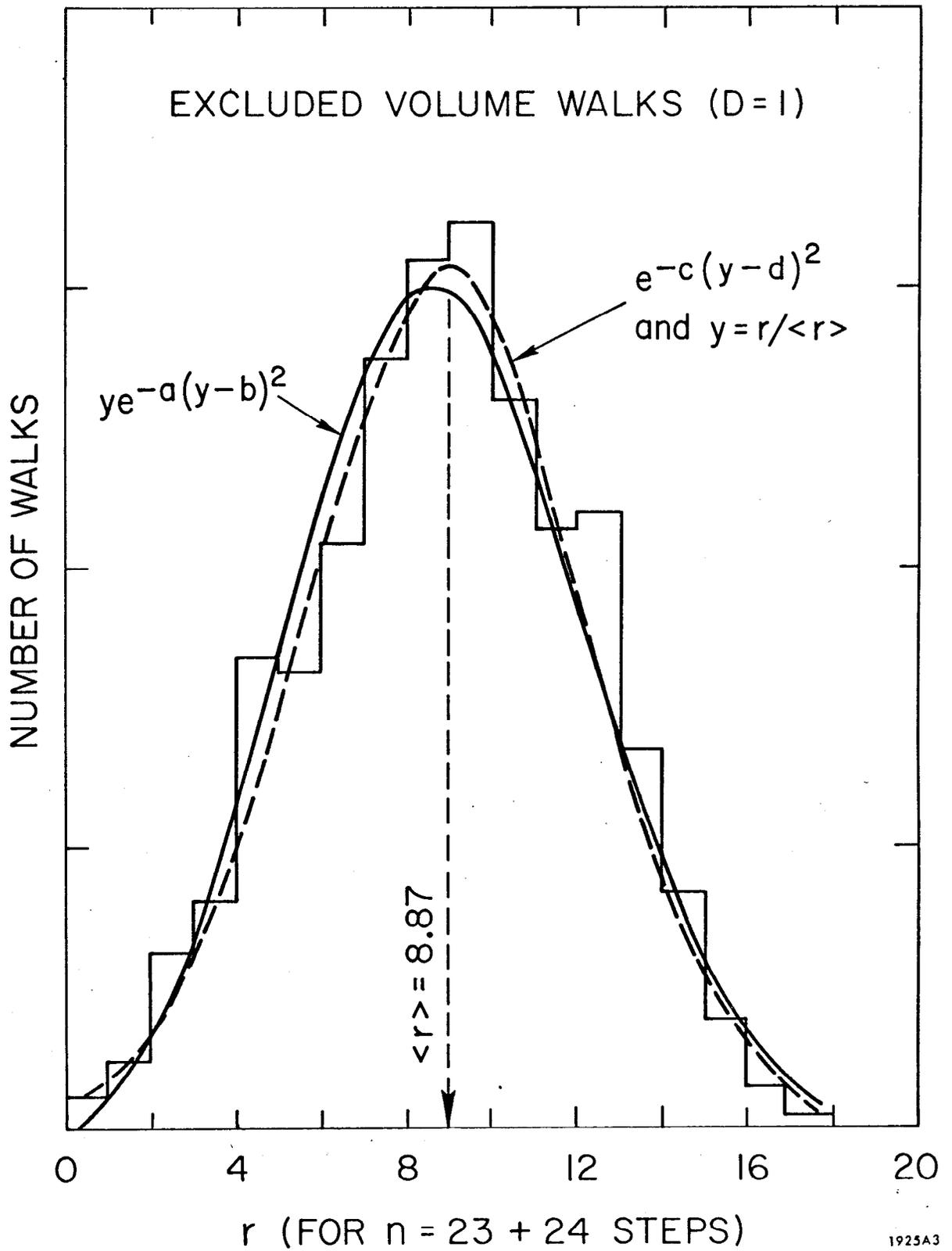


Fig 4

D=1 (EXCLUDED VOLUME) RANDOM WALKS

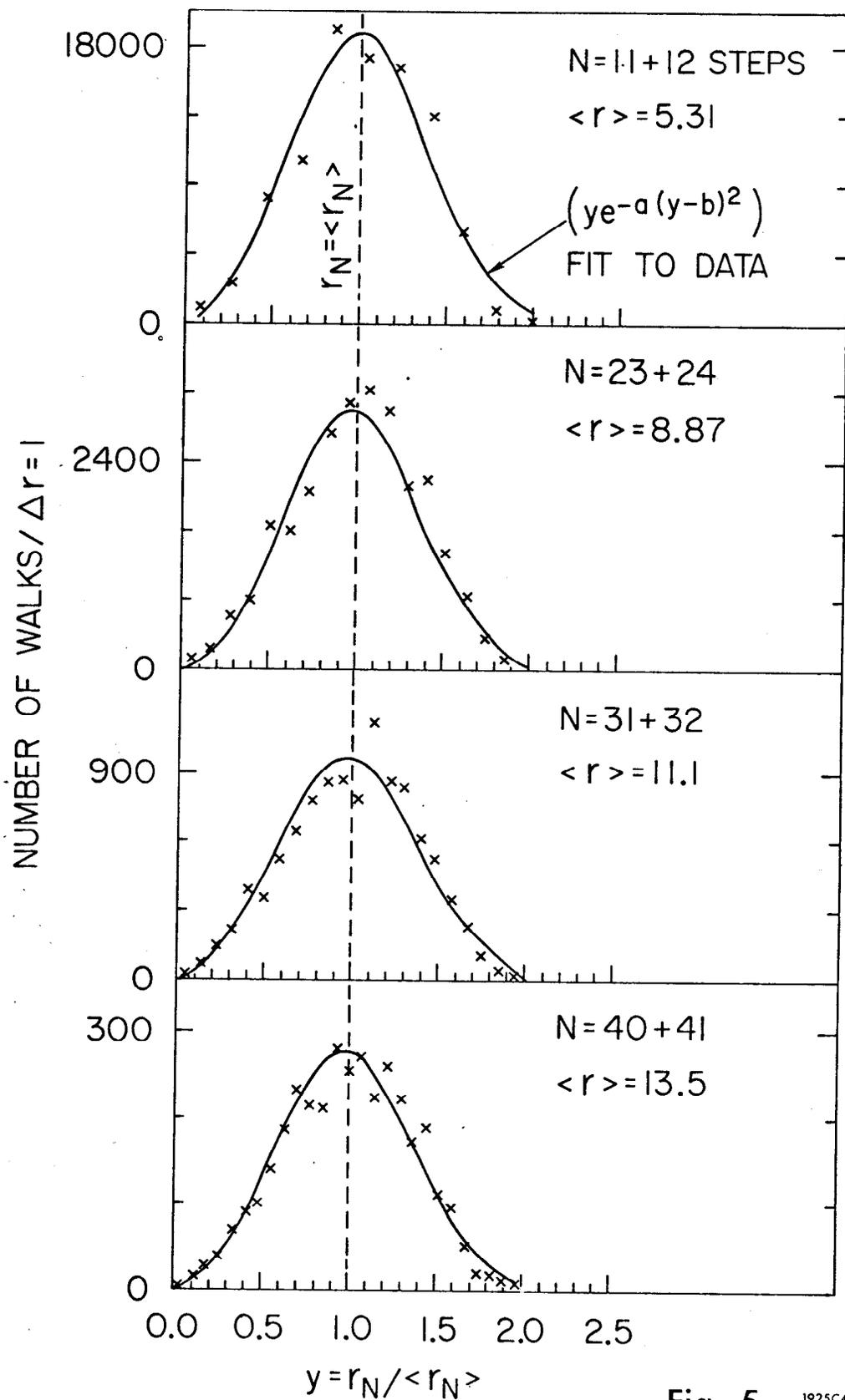


Fig. 5

1925C4

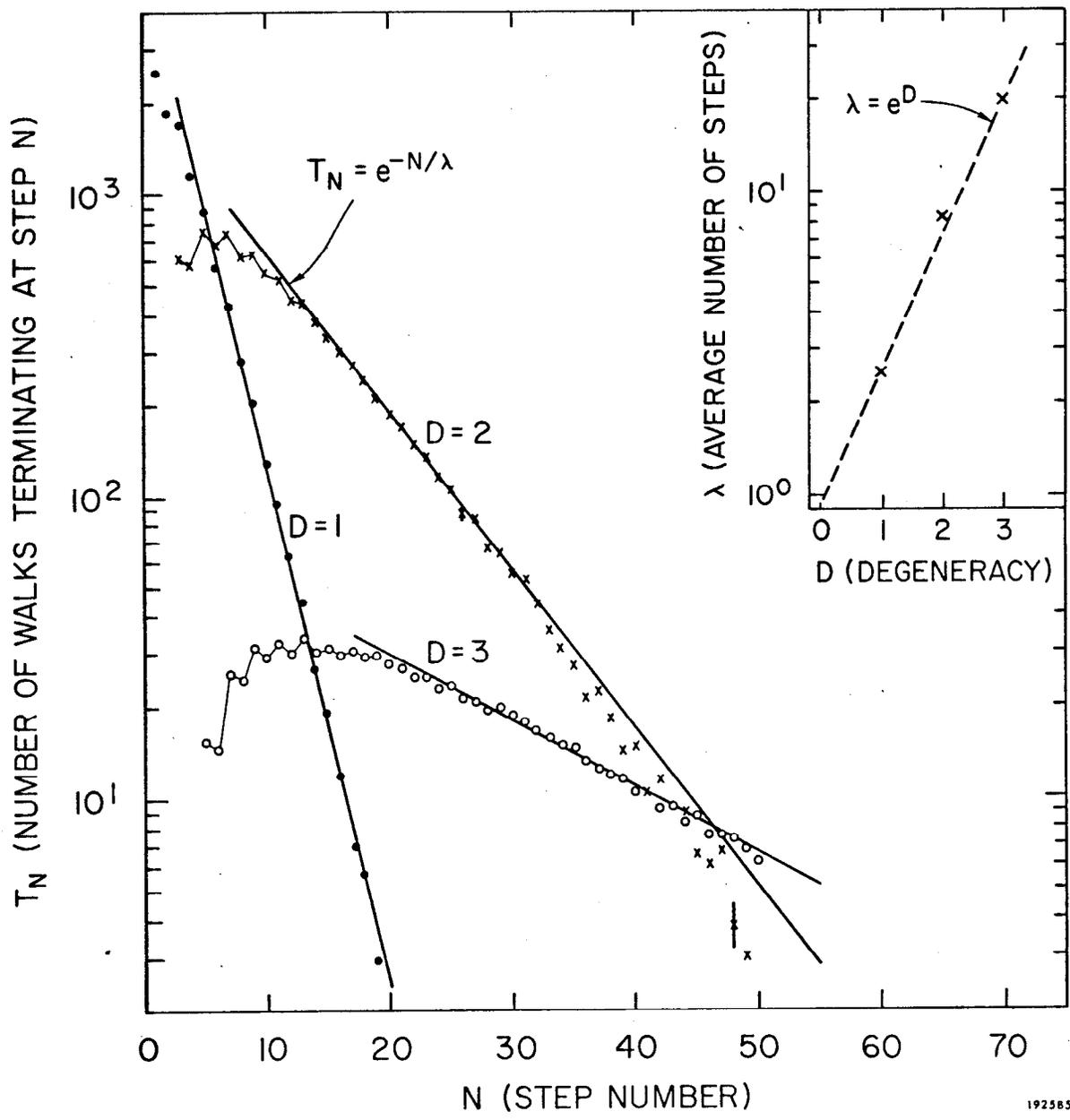


Fig. 6