THE BRITISH RUBBER PRODUCERS' RESEARCH ASSOCIATION

Publication No. 69

THE STATISTICAL LENGTH OF LONG-CHAIN MOLECULES.

By

L. R. G. TRELOAR.

Reprinted from the *Transactions of the Faraday Society*, No. 285,
Vol. XLII, Parts 1 and 2, January–February, 1946.

THE STATISTICAL LENGTH OF LONG-CHAIN MOLECULES.

By L. R. G. Treloar.

Received 29th June, 1945.

The problem of determining the statistical distribution of " displacement length " (*i.e.* end-to-end distance) of a long-chain molecule of given chain length is one of fundamental importance in the development of the theory of elasticity for rubber-like materials. Distribution functions for the paraffin [1] and polyisoprene [2] (natural rubber) chains have been published by the author in papers which will be referred to subsequently as I and II respectively. The method used depended on processes carried out graphically, no method having been discovered of treating the statistics of these molecules in precise analytical terms.

The simplest " molecular " structure which can be considered from the analytical standpoint is the chain of universally jointed equal links, and, as pointed out in I, an accurate solution to the problem of the statistical distribution of length for such a chain may be obtained when the number of links is small (*e.g.* ~ 6) by a method due to Rayleigh.[3] The method is impracticable when the number of links is large.

In the present paper a formula is derived which is applicable to much longer chains. This is obtained by a simple transformation of the result worked out by Hall [4] and Irwin [5] in 1927 in connection with the theory of random sampling. The relevance of the Hall-Irwin formula to the present problem was brought to the attention of the author by Dr. A. D. Booth of these laboratories.

The formula has been applied to chains of 25 and 100 links, and the resultant distribution curves are compared with those previously derived for the polyisoprene chain. The application of the results to the determination of the stress-strain relations of rubber will be considered in the following paper.

The x-Distribution for a Random Chain.

Let the chain be composed of n links, each of length l, such that each link is equally likely to be in any direction in space.* The distance between the ends of the chain is denoted by r, and the projection of r on a fixed line OX by x. It is required ultimately to find the distribution of r-values. As an intermediate step the distribution of x-values will be obtained.

[1] Treloar, *Proc. Physic. Soc.*, 1943, **55**, 345.
[2] Treloar, *Trans. Faraday Soc.*, 1944, **40**, 109.
[3] Rayleigh, *Phil. Mag.*, 1919, **37**, 321.
[4] Hall, *Biometrika*, 1927, **19**, 240. [5] Irwin, *ibid.*, 1927, **19**, 225.

* By this is meant that if the successive links are represented by vectors radiating from a point, the end points of these vectors, on the average, will be uniformly distributed over the surface of a sphere.

By eqn. (12) of I the probability that a single link has a component along OX between x and $x + dx$ is given by

$$p_1(x)dx = \frac{dx}{2l} \qquad (|x| < l), \quad . \quad . \quad . \quad (1)$$

i.e. the probability is constant over the range $-l$ to $+l$ and zero[1] elsewhere. A distribution of this type may be referred to as a "rectangular population." The x-component for the whole chain, which is the algebraical sum of the separate x-components for the individual links, is therefore the sum of n values of x chosen at random from the population represented by (1).

The problem of finding the x-distribution for the whole chain is therefore equivalent to finding the distribution of the mean of n samples chosen at random from a rectangular population. For the mean is given by

$$m = \frac{1}{n}\sum_i x_i$$

and in the present problem we require the total x-component for the chain which is

$$x = \sum_i x_i = nm.$$

The solution to this problem in statistics has been derived by Hall[3] and by Irwin[4] and is discussed briefly by Kendall.[5] The formula obtained by these authors gives the distribution of the mean of n samples from a rectangular population extending from $-\frac{1}{2}$ to $+\frac{1}{2}$ (*i.e.* corresponding to a link of length $\frac{1}{2}$), and is set out below:

$$f(m)dm = \frac{n^n}{(n-1)!}\sum_{s=0}^{k}(-1)^s\binom{n}{s}\left(m - \frac{s}{n}\right)^{n-1}dm \quad . \quad . \quad (2)$$

where k is defined by

$$\frac{k}{n} < m < \frac{k+1}{n}.$$

In this expression $f(m)$ is the probability that the mean has the value m, and $\binom{n}{s}$ represents the number of combinations of n things taken s at a time. It will be seen that the function $f(m)$ is represented by n arcs making contact at the successive points given by $m = k/n$. Both $f(m)$ and its first derivative are continuous functions. The function $f(m)$ is symmetrical about the ordinate $m = \frac{1}{2}$, and extends from $m = 0$ to $m = 1$.

To adapt this solution to the case of the chain it is only necessary to remember that what we require is the total x instead of the mean. The transformation is effected by putting $m = \frac{1}{2}(1 - x/nl)$, thus spreading the distribution out in the ratio $2nl$ along the axis of abscissæ and moving the origin to the position of the central maximum. At the same time the ordinates are reduced in the ratio $1/2nl$ so as to maintain unit area under the curve. The result is

$$p(x)dx = \frac{1}{2l}\frac{n^{n-1}}{(n-1)!}\sum_{s=0}^{k}(-1)^s\binom{n}{s}\left(m - \frac{s}{n}\right)^{n-1}dx \quad . \quad (3)$$

where $\frac{k}{n} < m < \frac{k+1}{n}$ and $m = \frac{1}{2}(1 - x/nl)$.

The r-Distribution Function.

To find the distribution of displacement length r we note that for chains of a given length r the corresponding distribution of x-values is given by an expression comparable with eqn. (1), *i.e.*

$$p_2(x)\mathrm{d}x = \frac{\mathrm{d}x}{2r} \qquad |x| < r$$

in which the limits of x are determined by r. In a continuous distribution of r-values, those chains whose r-values exceed a given (positive) value x will contribute to $p(x)$, whilst those for which $r < x$ make no contribution to $p(x)$. Hence the change in $p(x)$ over the range $\mathrm{d}x$ will be proportional to the number of chains whose r-values lie between x and $\mathrm{d}x$. In mathematical terms

$$-\left[\frac{\mathrm{d}p(x)}{\mathrm{d}x}\right]_{x=r} = \frac{1}{2r}P(r) \quad . \quad . \quad . \quad . \quad (4)$$

where $P(r)$ is the probability of a displacement length r. In this expression the positive half only of the symmetrical function $p(x)$ is considered, hence the factor $\frac{1}{2}$ is introduced on the right-hand side.

Application of (4) to the function (3) gives

$$P(r)\mathrm{d}r = \frac{r}{2l^2}\frac{n^{n-2}}{(n-2)\,!}\sum_{s=0}^{k}(-1)^s\binom{n}{s}\left(m - \frac{s}{n}\right)^{n-2}\mathrm{d}r \quad . \quad (5)$$

where
$$\frac{k}{n} < m < \frac{k+1}{n} \text{ and } m = \tfrac{1}{2}(1 - r/nl).$$

This is the exact distribution function covering the whole range of r. To illustrate the use of this formula, we may consider the case $n = 6$. We then have

$$P(r) = \frac{27r}{l^2}\sum_{s=0}^{k}(-1)^s\binom{6}{s}\left(m - \frac{s}{6}\right)^4$$

giving for the 3 separate regions of r

$$P(r) = \frac{27r}{l^2}\left(\frac{1}{2} - \frac{r}{12l}\right)^4 \qquad 4l < r < 6l,\ k = 0$$

$$P(r) = \frac{27r}{l^2}\left[\left(\frac{1}{2} - \frac{r}{12l}\right)^4 - 6\left(\frac{1}{3} - \frac{r}{12l}\right)^4\right] \qquad 2l < r < 4l,\ k = 1$$

$$P(r) = \frac{27r}{l^2}\left[\left(\frac{1}{2} - \frac{r}{12l}\right)^4 - 6\left(\frac{1}{3} - \frac{r}{12l}\right)^4 + 15\left(\frac{1}{6} - \frac{r}{12l}\right)^4\right] \qquad 0 < r < 2l,\ k = 2 \quad (6)$$

These expressions, when reduced, are identical with those derived by Rayleigh [3] by an entirely different method.

The functions $p(x)$ and $P(r)$ have been numerically evaluated for chains of 25 and 100 links each of unit length and the results are given in Tables I and II. For $n = 100$ the calculations were carried out by the Scientific Computing Service, for whose co-operation the author is indebted.

The form of the r-functions for $n = 25$ and $n = 100$ is shown in Fig. 1, in which $\log P(r)$ is plotted against $(r/r_m)^2$, $r_m(= nl)$ being the length of the fully-extended chain. Certain features of these functions are worthy of attention, namely :

(1) When r is small the curves approximate to the "normal" or Gaussian distribution, represented by the equation

$$P(r) = \frac{4\beta^3}{\pi^{\frac{1}{2}}} r^2 e^{-\beta^2 r^2} \quad (\beta^2 = 3/2nl^2) \qquad (7)$$

which is shown for comparison in Fig. 1.

(2) When r is large (*i.e.* $r > 0{\cdot}7r_m$) only the first term ($s = 0$), of the summation in (5) is important, so that

$$\frac{1}{r}P(r) \propto m^{n-2} \propto (1 - r/r_m)^{n-2}. \qquad (8)$$

(3) The general form of the function $P(r)$ is relatively insensitive to n, the number of chain-links.

TABLE I.

x AND r DISTRIBUTION.

x or r.	$\log_{10} p(x)$.	$-\log_{10} P(r)$.
For 25-link Random Chain.		
1	0·89	1·51
3	1·10	0·77
5	1·51	0·73
7	2·14	1·06
9	3·01	1·69
11	4·14	2·63
13	5·57	3·89
15	7·37	5·53
17	9·66	7·65
19	12·64	10·46
21	16·87	14·47
23	24·09	21·35
For 100-link Random Chain.		
2	1·187	1·806
6	1·395	1·061
10	1·811	1·034
14	2·437	1·366
18	3·279	1·984
22	4·342	2·867
26	5·631	4·006
30	7·155	5·400
34	8·925	7·054
38	10·955	8·978
42	13·261	11·187
46	15·864	13·70
50	18·79	16·54
54	22·07	19·74
58	25·76	23·34
62	29·90	27·40
66	34·57	31·99
70	38·89	37·22
74	46·01	43·26
78	53·18	50·33
82	61·80	58·85
86	72·61	69·52
90	87·07	83·82
94	109·04	105·54
98	156·27	152·28

Comparison with Paraffin and Polyisoprene Chains.

There is a close similarity between the distribution functions for the randomly-jointed chain of equal links and those previously obtained for the paraffin and polyisoprene structures. In Fig. 1, the distribution function for the 64-isoprene chain, represented by discrete points, is shown for comparison.

A more striking comparison is obtained by plotting $\log \frac{1}{r}P(r)$ against $\log (1 - r/r_m)$, which, when r is large, gives a straight line of slope $n - 2$ in the case of a random chain. (*Cf.* eqn. (8).) From Fig. 2 it is seen that this comparatively simple law holds approximately also for the polyisoprene chains. Assuming for the moment that the polyisoprene chain may be replaced by a random chain of a suitable number of links, then the equivalent number of links may be obtained from the slopes of these logarithmic plots. The results are given in Table II.

From the shape of the distribution at high r-values we are thus led to the conclusion that in this region, for all chain lengths, each isoprene unit is equivalent to about 1·44 random links.

Another way of obtaining the number of links in the equivalent random chain (if such there be) is to consider the exponent β^2 in the Gaussian approximation (7) for small r. Wall's [7] formula for the mean square length $\overline{r^2}$ of the polyisoprene chain leads to the value

$$\overline{r^2} = 3{\cdot}72Z \text{ (A.}^2\text{)},$$

[6] Kendall, *Advanced Statistics*, Vol. I (C. Griffin & Co.) 1943, p. 240.
[7] Wall, *J. Chem. Physics*, 1943, **11**, 67.

Z being the total number of bonds. For a 64-isoprene chain $Z = 256$, and since $\overline{r^2} = 3/2\beta^2$ we obtain for β^2 the value 0·00157.* The maximum length r_m for the 64-isoprene chain is 294 A.

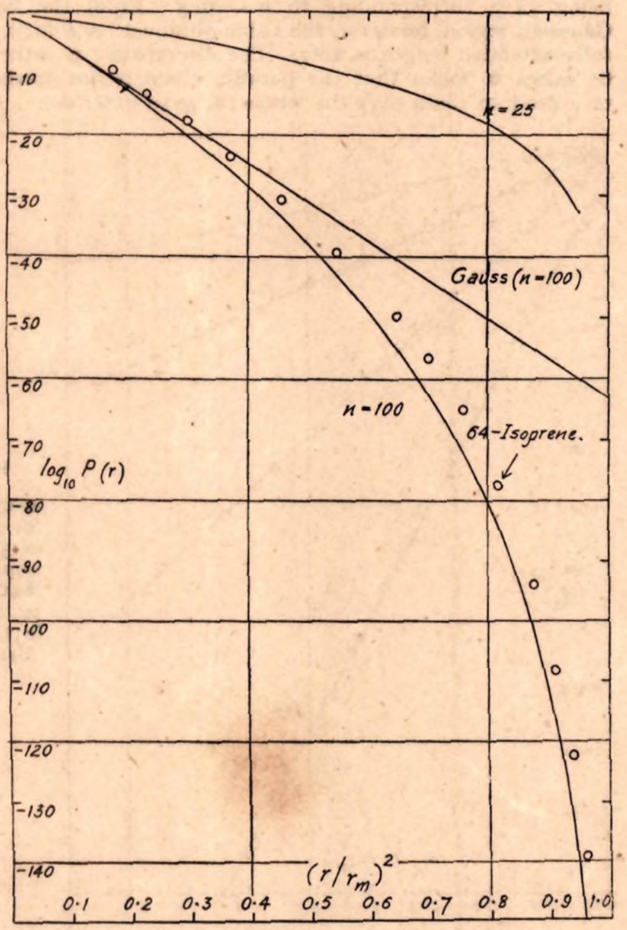

FIG. 1.

The distribution function for random chains of 25 and 100 links each of unit length. The function degenerates to the Gaussian form when r is small. The circles refer to the distribution for a polyisoprene chain of 64 isoprene units.

For a random chain of n links, each of length l, $\beta^2 = 3/2nl^2 = 3n/2r_m^2$. Hence, to find the equivalent number of links in the random chain having the same maximum length r_m as the 64-isoprene chain, we put $r_m = 294$ A. in the above expression. The result is $n = 90·5$, or 1·42 links per isoprene unit. This compares favourably with $n = 92·4$ obtained from the high-r region.

TABLE II.

No. of Isoprene Units in chain.	Slope $n-2$.	n.	n / No. of Isoprene Units.
32	44·0	46·0	1·437
64	90·4	92·4	1·443
128	184	186	1·452
256	369	371	1·448

We conclude therefore that *a long polyisoprene chain is statistically*

* This differs from the value 0·001325 given in II on account of the approximations involved in the previous method, which are discussed in the paper.

equivalent over the whole range of extension to a random chain containing about 1·42 links for each isoprene unit.

The paraffin chain is not quite so straightforward. The 80-link paraffin again gives a linear plot of log $P(r)/r$ against log $(1 - r/r_m)$ the slope being 34·5, corresponding to $n = 36{\cdot}5$. From the value of β^2 in the Gaussian region, however, the value obtained for n for a chain of the same fully-extended length is 26·7. The discrepancy is rather large, and must be taken to mean that the paraffin chain is not statistically equivalent to a random chain over the whole range of extension.

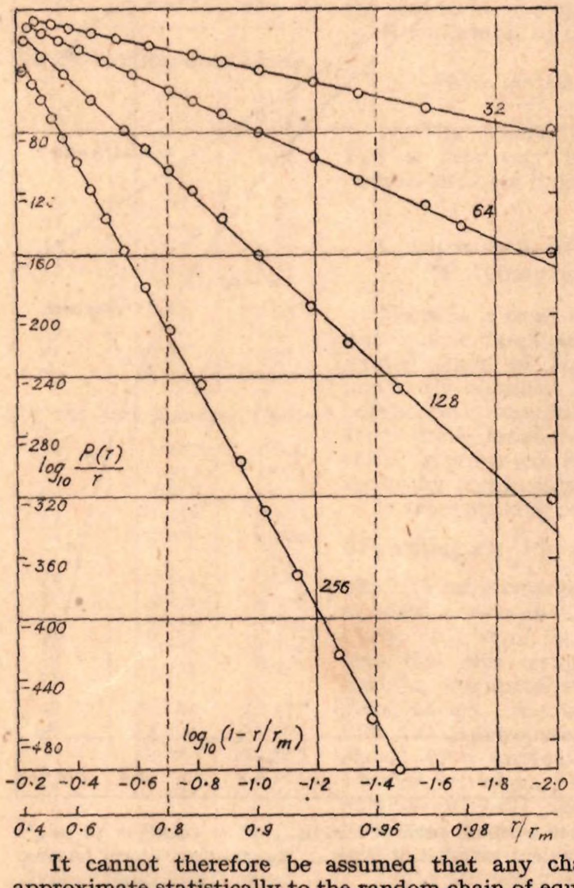

FIG. 2.

Plot of $\log \frac{1}{r} P(r)$ against log $(1 - r/r_m)$ for polyisoprene chains of 32, 64, 128 and 256 isoprene units, showing that $P(r)$ approximates to the same form as in the case of random chains.

It cannot therefore be assumed that any chain-like structure will approximate statistically to the random chain of equal links over the whole range of r. The fact of this approximation in the case of the polyisoprene chain is a result of its particular geometrical structure.

Summary.

A formula is derived for the complete function representing the probability of a given distance between the ends of a chain of universally jointed equal links. The formula is computed for chains of 25 and 100 links. The distribution functions derived from this formula are compared with those previously worked out by an independent method for polyisoprene and paraffin chains. It is shown that the polyisoprene chain is statistically equivalent to a randomly-jointed chain of length corresponding to 1·42 links per isoprene unit.

This work forms part of a programme of fundamental research on rubber undertaken by the Board of the British Rubber Producers' Research Association.

PRINTED IN GREAT BRITAIN BY THE UNIVERSITY PRESS, ABERDEEN