

Statistics of branched polymers

This article has been downloaded from IOPscience. Please scroll down to see the full text article.

1990 J. Phys. A: Math. Gen. 23 4055

(<http://iopscience.iop.org/0305-4470/23/17/033>)

View [the table of contents for this issue](#), or go to the [journal homepage](#) for more

Download details:

IP Address: 129.252.86.83

The article was downloaded on 01/06/2010 at 08:56

Please note that [terms and conditions apply](#).

COMMENT

Statistics of branched polymers

M J Godfrey

Department of Physics, University of Warwick, Coventry CV4 7AL, UK

Received 4 April 1990

Abstract. I consider the number of distinct non-cyclic monohaloalkanes, compared with the number of configurations of these molecules on a Bethe lattice. It is found that a finite fraction of the bonds in a large molecule fail to generate fresh configurations when the attached branch is rotated.

The solution of structural combinatorial problems is important in the study of a number of statistical models in physics. For example, in the theory of polymer solutions, calculating the number of configurations of a chain molecule is a first step toward finding its free energy, which is mainly entropic in origin. In studies of gelation, parameters of interest are the gel fraction and the distribution of sizes and structures of gel fragments above and below the gelation threshold (de Gennes 1979). In an early treatment, Flory (1953) modelled the gel transition by what amounted to a percolation problem on a Bethe lattice. More generally in percolation problems, the variation with concentration of the distribution of sizes of subpercolation clusters usefully characterizes the percolation transition.

We can view the subpercolation clusters in a way which may be of interest in assessing the validity of percolation theory as a model of gelation. We regard the set of lattice animals as different conformations of an irreducible underlying set of 'lattice molecules'. In other words, we view as identical those animals that can be deformed into one another by a specified set of elementary limb manipulations (specifically excluding removal of a limb). For example, we might regard the two distinct 3-site animals on a square lattice—an L-shape and an I-shape—as different conformations of the same linear molecule of three atoms. Proceeding in this way, starting from a distribution of clusters we would 'quotient out' the configurations of the underlying set of molecules to leave us with the distribution of distinct molecules. This kind of approach might be of use in assessing the yield of different products in a theoretical study of polymerization or vulcanization modelled by the corresponding problems of reacting monomers or polymers on a lattice.

Let us focus on one feature of the outlined scheme, and consider the calculation of the number of structural isomers of a molecule of known chemical composition. In this case the local configuration of bonds issuing from an atom is given, and the group of configurations of the whole molecule is generated by rotations around the bonds connecting the atoms. As a simple model problem of this kind, consider the chemistry of non-cyclic saturated hydrocarbons; in particular, the molecules with chemical formula $C_nH_{2n+1}X$, where X and H are monovalent atoms, and C is tetravalent with bonds that project towards the vertices of a circumscribed tetrahedron. We will see later that the atom X, by providing a reference point on the molecule, facilitates the

finding of a recurrence relation between the numbers of molecules with different numbers of atoms.

In addition to ignoring the complications of loop structures in the molecules, we shall further neglect interactions between the different branches of the molecule; we assume that it is always possible to manipulate a molecule into any configuration that respects the local arrangement of bonds at a carbon atom. This will involve treating the branches as ghosts. Since we are interested only in the *numbers* of molecules with n carbon atoms, the last assumption can be incorporated by imagining the molecules to lie on a Bethe lattice; on identifying the atom X with a halogen, we may state our problem as that of enumerating the 'Bethe haloalkanes' and their configurations.

In generating the different configurations of a molecule, rotation by an angle $2\pi/3$ about a given bond $\text{XR}-\text{CR}_1\text{R}_2\text{R}_3$ (where R_1 , R_2 and R_3 denote saturated alkyl branches) is equivalent to cyclic interchange of the outlying three branches. Since R_1 , R_2 and R_3 are different in general, we might expect—in the limit of large molecules—that rotation about a given bond would generate on average three (or very nearly three) configurations of the molecule on the lattice. We will see eventually that this supposition is very far from being correct.

We build up trees from a common root, denoting by g_n the number of trees with n carbon atoms ('nodes'), and defining $g_0 = 1$. A tree with $n + 1$ nodes can be constructed in two stages. First we attach a single node to the root, leaving three 'dangling bonds'. We then attach to these bonds three sub-trees ('branches') containing i , j and k nodes, with $i + j + k = n$. If n is not exactly divisible by 3, at least two of the branches must be different, so that there are three distinct configurations, (i, j, k) , (j, k, i) and (k, i, j) , which represent essentially the same tree. When n is a multiple of 3 the situation is similar, except for the $g_{n/3}$ configurations with three identical branches: each of these configurations corresponds to precisely one tree.

Let us call the set of unsymmetrical trees U and the set of symmetrical trees S . Then we can write

$$g_{n+1} = \frac{1}{3} \sum_U 1 + \sum_S 1 = \frac{1}{3} \sum_{U \cup S} 1 + \frac{2}{3} \sum_S 1$$

where the first sum is taken over *all* the trees we have built up, without regard to their symmetry. Taking into account in this way the difference between the symmetrical and the unsymmetrical trees, we find the following expressions for the number of distinct trees with $n + 1$ nodes,

$$g_{n+1} = \begin{cases} \frac{1}{3} \sum_{i=0}^n \sum_{j=0}^n g_i g_j g_{n-i-j} + \frac{2}{3} g_{n/3} & \text{when } n \text{ is divisible by 3} \\ \frac{1}{3} \sum_{i=0}^n \sum_{j=0}^n g_i g_j g_{n-i-j} & \text{otherwise.} \end{cases} \quad (1)$$

A list of values of g_n is given in table 1.

To assist in the asymptotic solution of (1) we combine the two parts by introducing a generating function

$$G(t) = \sum_{n=0}^{\infty} g_n t^n \quad (2)$$

where t is a formal parameter. The asymptotic behaviour of g_n for $n \rightarrow \infty$ is closely related to the singular behaviour of G regarded as a function of the complex variable t .

Table 1. Numbers of distinct monohalocarbons, g_n , containing n carbon atoms. For large n , g_n should vary according to (8), which, for comparison with the values in the table, gives $g_{20} \approx 8.40 \times 10^7$ and $g_{50} \approx 6.79 \times 10^{22}$.

| n | g_n |
|-----|-----------------------|
| 0 | 1 |
| 1 | 1 |
| 2 | 1 |
| 3 | 2 |
| 4 | 5 |
| 5 | 11 |
| 6 | 28 |
| 20 | 8.23×10^7 |
| 50 | 6.73×10^{22} |

Equation (1) is equivalent to the following functional equation for $G(t)$:

$$G(t) = 1 + \frac{1}{3}tG(t)^3 + \frac{2}{3}tG(t^3) \tag{3}$$

as we can see by expanding each side in powers of t and equating coefficients of t^n . Now the coefficients g_n are all real and positive, so that G increases monotonically to its first singularity at $t = r$ on the real axis as t increases from 0. The number r is the limiting value of the ratio g_n/g_{n+1} , $n \rightarrow \infty$, and so must be less than unity. This leads to a convenient simplification of (3) near the singularity: we have $r^3 < r$, so that $G(t^3)$ is a regular function of t at $t = r$. Hence (3) can be regarded as a cubic equation for G with coefficients that are regular in t at r . We rewrite (3) as

$$G^3 - 3CG + 2D = 0$$

with $C = 1/t$ and $D = (G(t^3) + 3/2t)$. The solutions of this cubic equation take the form (Abramowitz and Stegun 1972)

$$G = [-D + \Delta]^{1/3} + [-D - \Delta]^{1/3} \tag{4}$$

where $\Delta^2 = (D^2 - C^3)$. Since $C = 1/t$ is non-zero for finite t , neither cube root in (4) vanishes. The singularity in G is therefore due to the vanishing of the discriminant Δ^2 ,

$$\Delta^2 = (G(r^3) + 3/2r)^2 - 1/r^3 = 0 \tag{5}$$

which is readily solved for r to any desired degree of accuracy, since $G(r^3)$ converges rapidly. We could use (3) and (5) to eliminate $G(r^3)$ in favour of $G(r^9)$, which is still more rapidly converging, but there is little practical advantage to be gained from this. From (5) and the first few values of g_n in table 1 we obtain $r = 0.304\ 218\ 41$.

The cube roots in (4) can be chosen so that the solution branch increases with t from $t = 0$, corresponding to our original series representation (2). In the neighbourhood of r , $G(t)$ has the form

$$G(t) = 1/\sqrt{r} - B(r-t)^{1/2} + O(r-t) \quad t \rightarrow r \tag{6}$$

where

$$B^2 = 1/r^2 - 1/r^{3/2} + 2r^{5/2}G'(r^3) \tag{7}$$

and $G'(t)$ denotes the formal derivative of $G(t)$ with respect to t . (Naturally, once we know the form of the singularity in G we can try a series solution for (3) of the form (6); in practice a systematic expansion of this kind would be useful if higher order corrections were required to the asymptotic form of g_n to be calculated below.)

The limiting behaviour (6) of $G(t)$ allows us to find the asymptotic form of the coefficients g_n . We expand $(r-t)^{1/2}$ in a power series in t and compute the general coefficient. This gives

$$g_n \sim B\sqrt{r} \frac{1 \cdot 3 \cdot 5 \dots (2n-3)}{n!} \frac{1}{(2r)^n}$$

$$\sim \frac{B}{2\sqrt{\pi}} r^{(1/2)-n} n^{-3/2}.$$

The coefficient B can be calculated directly from (7), giving

$$g_n \sim DQ^n/n^{3/2} \quad n \rightarrow \infty \quad (8)$$

where $D=0.346\ 304\ 27$ and $Q=1/r=3.287\ 1120$.

(The reasoning presented above can easily be repeated for the case of trivalent atoms in which the three bonds extending from each atom lie in one plane. The numbers of molecules g_n takes the same form (8) with $D=0.791\ 6032$ and $Q=1/r=2.483\ 2536$.)

Now that we have the asymptotic form for g_n , we might, as suggested above, hope to relate it very simply to the number of trees—let us call it h_n —with n carbon atoms and non-rotatable bonds. There are n bonds to rotate into each of three different configurations, so that to within factors of order unity we might expect $h_n \sim 3^n g_n$. This, however, turns out to be an exponentially large overestimate, as we shall see by calculating the h_n explicitly.

The generating function $H(t) = \sum_{n=0}^{\infty} h_n t^n$ satisfies the following equation, analogous to (3) above,

$$H(t) = 1 + tH(t)^3.$$

Writing $H = 1 + K(t)$ we have $t(K) = K/(1+K)^3$, which is in a form suitable for application of Lagrange's formula for reversion of power series (Whittaker and Watson 1927). We find

$$h_n = \frac{(3n)!}{n!(2n+1)!} \sim \frac{1}{4} \sqrt{\frac{3}{\pi}} \frac{(27/4)^n}{n^{3/2}}$$

$$\propto 3^n \cdot (9/4)^n / n^{3/2} \quad n \rightarrow \infty.$$

(In the trivalent case this becomes $h_n \sim \pi^{-1/2} 2^n \cdot 2^n / n^{3/2}$, to be compared with $2^n g_n$.) This shows that $3^n g_n$ overestimates h_n by a factor $(Q/(9/4))^n \approx 1.46^n$ even when n is small enough (say, 10) for the neglect of interference of sub-branches to be a reasonable approximation.

The reason for this is to be sought in (1). For large n the $g_{n/3}$ 'correction term' on the right-hand side is negligible, being only one term among $O(n^2)$ others, all of which are much larger than it. In fact it is only for *small* n that the correction has any appreciable effect; but these lower g_n 's still determine the asymptotic form of g_n through the *first* term of (1).

We can express this result in a much more concrete way. The discrepancy is due to 'lazy' bonds attached to 3-fold symmetric chain-terminating groups such as CH_3 : rotation about these bonds does not generate new tree configurations. The exponential error shows that the number of lazy bonds is of order n ; and the fraction of the total number of bonds for tetravalent [trivalent] carbon is $\log Q/\log(9/4) \approx 0.345$ (or $\log Q/\log 2 \approx 0.312$). It is perhaps surprising that the fraction of lazy bonds is not, as far as I can see, a number that can be expressed simply in terms of integers.

Acknowledgments

The author would like to thank W D Brown, J M F Gunn and M W Long for a number of useful discussions and comments. He also thanks the SERC for financial support.

References

- Abramowitz M and Stegun I A 1972 *Handbook of Mathematical Functions* (New York: Dover)
Flory P J 1953 *Principles of Polymer Chemistry* (Ithaca, NY: Cornell University Press)
de Gennes P G 1979 *Scaling Concepts in Polymer Physics* (Ithaca, NY: Cornell University Press)
Whittaker E T and Watson G N 1927 *A Course of Modern Analysis* (Cambridge: Cambridge University Press)