THE
PHYSICAL REVIEW

A journal of experimental and theoretical physics established by E. L. Nichols in 1893

SECOND SERIES, VOL. 103, NO. 1 JULY 1, 1956

Combinatorial Problems Suggested by the Statistical Mechanics of Domains and of Rubber-Like Molecules

H. N. V. TEMPERLEY*
National Bureau of Standards, Washington, D. C.
(Received August 9, 1955)

The mathematical problem of enumerating the number of connected domains that can be drawn on a plane square lattice is studied by several methods, and results that are believed to be very nearly correct for large domains are obtained, while a slightly modified version of the problem can be solved in closed form. Comparison of the various methods tried leads to conclusions which, besides their purely mathematical interest, have a bearing on a large number of physical problems, such as ferromagnetism, order-disorder in alloys, the theory of solutions, fusion and evaporation, the configuration of polymer molecules and gel formation. The comparison of the various methods also gives information on what may be expected of an approximate theory of a phase-transition.

1. INTRODUCTION

THE writer has pointed out elsewhere[1,2] that there is a close relationship between the theory of partition of numbers and a certain type of problem in statistical mechanics, the computation of the number of ways in which a given total quantity of energy can be distributed among "particles" occupying a given set of energy levels, subject to various auxiliary conditions. Usually, the total number of "particles" is considered to be fixed. Wannier[3] and Kac and Ward[4] have called attention to the relationship between models of the Ising type and the number of diagrams, consisting entirely of closed polygons, that can be drawn on lattices of various kinds, e.g., the two-dimensional square lattice, subject to no line being used twice and to the total perimeter of all the polygons being specified. Various other writers have studied generalizations of the "random-walk" problem in connection with the theory of high polymers and the properties of rubber-like materials and proteins. Here the objective is to enumerate the number of possible "random walks" in which, however, various constraints are assumed, for example paths that cross themselves may be excluded, the possible angles between two successive "steps" may be prescribed, or one may allow only paths all of whose "steps" lie on a specified lattice. The introduction of such constraints vastly complicates the solution of such problems, but the study of their effects is, nevertheless, of considerable interest. The possible number of ways in which, e.g., a rubber molecule can exhibit itself, is severely limited by steric considerations, and it would be of interest to have information on the "mathematical" effect of applying such constraints, as opposed to the mere collection of numerical and asymptotic results.

In this paper, we study a set of problems closely allied both to the "semirandom" walk and to the Ising problem. The type of problem we study is the number of ways in which a single domain, of given perimeter and total area, can be marked off on a plane square lattice. In its full force, this problem has proved extremely stubborn, though a partial solution of it can be given, and a slight modification of it can be solved with ease. Other ways in which problems of this type can be expressed are, e.g., the enumeration of "Cook's tours" of specified types on the lattice, or the enumeration of the ways in which n patience cards can be set out in a tableau, each card having contact with at least one other.

The "domain" problem may be regarded as a special case of the "walk" problem, namely possible "domain"

* Present address: Riversdale, Grantchester, Cambridge, England.
[1] H. N. V. Temperley, Proc. Roy. Soc. (London) **A199**, 361 (1949).
[2] H. N. V. Temperley, Proc. Cambridge Phil. Soc. **48**, 683 (1952).
[3] G. H. Wannier, Revs. Modern Phys. **17**, 50 (1945).
[4] M. Kac and J. C. Ward, Phys. Rev. **88**, 1333 (1952).

boundaries consist of "walks" that return precisely to their starting point.

2. PHYSICAL SITUATIONS GIVING RISE TO THE DOMAIN PROBLEM

There has been considerable controversy in recent years on the exact state of affairs near the critical point of a liquid,[5] one of the points at issue being the question of whether there really exists a finite region of temperature for which the surface tension vanishes and yet there is still a measurable difference of density between the two phases. A very important advance on the experimental side was made by Schneider and Atack[6] who seem to have shown quite definitely that, near the critical region, extremely sharp density gradients can be produced by the effect of gravity, so that many of the older experiments, which have been interpreted as proving the existence of a two-density region even in the absence of a meniscus, might have given different results if the effect of gravity could have been eliminated. Ordinarily, the effect of gravity on the density of a liquid is negligibly small, but it may be necessary to take it into consideration near the critical temperature, where the compressibility becomes very large.

Mayer and Mayer[7] have attempted, on theoretical grounds, to argue that the temperature at which surface tension vanishes *may* be distinct from that at which the densities become equal, but no actual *model* showing such an effect has yet been found.

Accordingly, it is of some interest to study models which are sufficiently realistic to exhibit properties analogous to the condensation of a gas and are at the same time simple enough to enable their analytic consequences to be worked out exactly. Yang and Lee[8] have called attention to the fact that the Ising model of a ferromagnetic can be reinterpreted to describe the "lattice gas." The vessel is supposed to be divided into cells, and we distinguish only between configurations in which a given cell is empty and those in which it is occupied by a molecule. The "vapor" phase is supposed to consist of single molecules and small domains, while the liquid phase contains large domains. By a domain we mean an aggregate of molecules in cells which are in contact with one another. This is a very crude representation of a real gas, in that the attractions and repulsions between molecules are only described very roughly, but the model does show properties strongly resembling the liquefaction of a real gas. Among other results Yang and Lee[8] have proved that only *one* transition temperature is associated with the "liquefaction"—and this last result apparently holds for a large class of models besides the "nearest neighbor interaction" one that they consider explicitly. Since Mayer and Mayer[7] postulate *two* singularities in the partition function of a real imperfect gas, the lower transition temperature being associated with the vanishing of the surface energy between phases and the phases remaining distinguishable up to the higher transition temperature, it is of interest to compute the "boundary free-energy" as a function of temperature for the simplified models considered by Yang and Lee.[8] In the remainder of the paper we consider the two-dimensional square lattice with interactions between nearest neighbors only. The calculation of the boundary free energy is equivalent to the enumeration, in the magnetic version of the model, of all possible boundaries separating a connected domain of "positive" magnets from a similar domain of "negative" magnets. We wish to decide whether or not the free energy associated with the formation of a domain boundary can vanish at any temperature below the bulk transition temperature.

A simple argument of Onsager's, quoted by Temperley,[2] seems to show that, for the plane square lattice, the boundary free energy remains finite at all temperatures below the bulk transition temperature. Onsager showed that the free energy associated with the formation of a domain boundary, *restricted to be of a certain type*, vanished at precisely the "bulk" transition temperature. This argument can, however, be criticized on the ground that the type of path considered by Onsager is *not* the most general type of path that can form a definite domain boundary. It remains, in principle, possible that the constraints introduced by Onsager are too restrictive, and result in the rejection of too many of the paths that can serve as domain boundaries. If so, a relaxation of these constraints would lead to a significant increase in the boundary entropy, which, in turn, would mean that the boundary free energy would vanish at a temperature significantly *below* the bulk critical temperature (instead of vanishing at exactly that temperature, as Onsager calculates). We should then have a model that was behaving in exactly the way envisaged by Mayer and Mayer.[7] However, the present investigation seems to establish that, in the lattice gas model, the boundary tension remains finite right up to the "bulk" critical temperature.

3. DEFINITION OF A GENERATING FUNCTION, AND ITS POSSIBLE PHYSICAL INTERPRETATIONS

Any particular problem of this class may be held to be completely solved if we have an expression for the corresponding generating function. This function is such that there is a 1–1 correspondence between each term and each possible path. As a simple example, consider the random-walk problem, in which each step is of fixed length and is constrained to lie upon a two-dimensional square lattice. There are then four possible types of step (up, down, to the right, and to the left) which we associate respectively with the factors $y\epsilon$, y/ϵ, $x\delta$, x/δ

[5] See, for example, the Report of the Paris Conference on Phase Changes, (1952).
[6] W. G. Schneider and G. Atack, J. Phys. Chem. **55**, 532 (1951).
[7] J. E. Mayer and M. G. Mayer, *Statistical Mechanics* (John Wiley and Sons, Inc., New York, 1940).
[8] C. N. Yang and T. D. Lee, Phys. Rev. **87**, 404, 410 (1952).

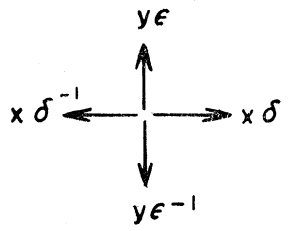

FIG. 1. The four types of step and the corresponding factors.

(Fig. 1). The function describing all possible numbers of steps is then

$$[1-(x\delta+x\delta^{-1}+y\epsilon+y\epsilon^{-1})]^{-1}, \tag{1}$$

because, if we expand the nth power of the expression in round brackets, each term in the expansion corresponds to the selection of just one term from each of the n factors, that is to just one possible choice of n "legs." The index of x corresponds to the total *number* of horizontal "legs," the index of δ to the total horizontal *displacement* between the two ends of this particular path. The various restricted types of path and domain that we shall study will have generating functions that could be obtained by expanding function (1), and then deleting all terms that describe paths which do not conform with the restrictions.

Let us consider the one-dimensional version of expression (1), (putting $y=0$). Its expansion may be written

$$\phi(x,\delta)=\sum_n \sum_m p_{n,m}x^n\delta^m,$$

where $p_{n,m}$ represents the number of different paths made up of n legs, whose beginning and end are m steps apart. If, as here, ϕ is a known function, it may be possible to determine an algebraic expression for $p_{n,m}$ from it. (In more complicated cases, it may be necessary to resort to the method of steepest descents or some similar technique). For many purposes, it is convenient to work with the generating function itself rather than with the individual p's. Suppose, as a simple example, that we wish to determine the average value of m^2 for all paths of fixed total length n. This may be written

$$\langle m^2\rangle_{\text{Av}}=\sum_m m^2p_{n,m}/\sum_m p_{n,m},$$

which is equal to the ratio of the coefficient of x^n in $[\delta(\partial/\partial\delta)]^2\phi$ to the coefficient of x^n in ϕ, the variable δ being set equal to unity after the differentiations. It is very often easier to obtain results of this type from the function ϕ itself rather than to attempt summations involving the $p_{n,m}$'s directly, the function ϕ acting as a sort of "packaged" form of the $p_{n,m}$'s. It is this fact that makes the generating function technique a valuable one, and it was much used by Cayley, Sylvester, and MacMahon in their work on partition of numbers and similar problems, and also by Mayer and Mayer[7] in their investigations on the statistical mechanics of an imperfect gas and related problems in statistical mechanics. As another simple example, we may want to study just those paths that return exactly to their starting point. The generating function for these special paths is simply

$$\sum_n p_{n,0}x^n=\text{terms independent of } \delta \text{ in } \phi(x,\delta).$$

We can carry out similar operations for paths in two and three dimensions, if we have a compact expression for the generating function, and we now consider some of these.

(a) We may study the variation, with n the number of steps, of the mean path "spread" (distance between beginning and end of path) and the distribution of path spreads about this mean value. For any given term in the expansion of (1), the square of the path spread can be obtained by applying the operator $[\delta(\partial/\partial\delta)]^2+[\epsilon(\partial/\partial\epsilon)]^2$. It is well known that, for generating function (1), the mean spread is proportional to $n^{\frac{1}{2}}$ for large n, and that the distribution about this mean approaches the normal type. It has been one major objective to find out whether such general results still hold if *restrictions* are placed on the types of permitted path, restrictions of the kind imposed in practice on rubber-like molecules. See Wall[9] and Wall *et al.*[10]

(b) We may select from the expansion of a generating function such as (1) only those terms corresponding to *closed* paths—that is, the terms in the expanded function independent of both δ and ϵ. This will usually be done by replacing δ and ϵ by the imaginary exponentials $e^{i\omega_1}$, $e^{i\omega_2}$ and then integrating with respect to ω_1, and ω_2. We call this operation D. The indices of x and y keep a record of the *perimeter* of each domain, and we sometimes also want to study domains enclosing different areas of the lattice, our problem then becoming a "lattice-point" one. To take care of this further requirement, we should have to take a "path" function such as (1), perform operation D on it, examine each closed path in turn, and multiply the corresponding term in the generating function by z^m, where m is the number of squares enclosed by the path.

In practice, it may frequently happen that we are able to calculate the "path" function but not the *complete* "domain" generating function complete with z factor, or, again, we may possess the "domain" function and yet be unable to calculate the function generating "open" paths. It is usually much more difficult to derive a "domain" function if we want a record of both the perimeter and the enclosed area.

In the application of such results to statistical mechanics, we have always to consider two types of physical situations associated with any one of the "selector" variables x, y, and z.

(a) We may have a situation in which we are only interested in terms with a *fixed index*, e.g., we might have the problem of the number of ways in which a fixed *number of squares* can be arranged to form a connected domain. This form of problem occurs when we are

[9] F. T. Wall, J. Chem. Phys. **21**, 1914 (1953).
[10] Wall, Hiller, and Wheeler, J. Chem. Phys. **22**, 1036 (1954); Wall, Hiller, and Atchison, J. Chem. Phys. **23**, 2314 (1955).

applying the Ising model to the theory of solutions or order-disorder problems, or to the theory of an imperfect gas when we prescribe the number of atoms in an enclosure. The corresponding mathematical operation is to select a particular power of z from the domain generating function.

Again, we might ask for the number of ways in which we can, subject to certain constraints, form a path consisting of m horizontal steps and n vertical steps. For this we should select the coefficient of $x^m y^n$ from the appropriate generating function. There is an obvious extension of this to the case in which m and n are not specified separately, but only their sum $m+n$. This corresponds to the problem of enumerating paths consisting of a specified total number of "steps," or domains of a specified perimeter.

(b) We may have a situation, e.g., in the ferromagnetic or adsorption problem, in which we are interested in *all* powers of one of the variables such as z. In such problems, instead of prescribing a particular power of z, we ascribe a definite numerical value to this variable, and then seek to evaluate the generating function and its derivatives. Assigning a definite value to z would assign different *probabilities* to domains of different sizes, thus taking account, e.g., of the effect of an applied magnetic field. We might be interested in, e.g., the average size of a domain or the mean total length of its boundary. (We have to consider *all* powers of our selector variable in problems of this type, because, in the magnetic case, we have to allow for the elementary magnets "flipping over" from one orientation to another, while, in the adsorption case, it is customary to consider an array of adsorption sites in equilibrium with a gas phase. The probability that a given site is occupied or empty depends on the gas pressure, but all configurations from completely full to completely empty are, in principle, possible.)

Such problems are solved if we can calculate various derivatives of the generating function as a function of the selector variable. For a full discussion of the relationship between these two types of problem, corresponding respectively to the "ordinary" and "grand" partition functions, the reader is referred to Rushbrooke.[11] In practice, we shall be mainly concerned with this second type of application. In the ferromagnetic problem we shall be interested in locating the values of the variables for which the mean boundary length of a domain begins to increase very rapidly, that is, at which the generating function, considered as a function of the selector variables, first becomes singular. This will correspond to the possibility of large "reversed" domains appearing in the spontaneously magnetized ferromagnetic, to the breakdown of long-range order in an alloy, and to the possible occurrence of large vapor bubbles in the liquid phase at no cost in free energy, which we interpret as the vanishing of surface tension.

[11] G. S. Rushbrooke, *Statistical Mechanics* (Oxford University Press, Oxford, 1949).

In other words, we ask whether the partition function associated with a single connected domain becomes singular for the same values of the variables as does the more general partition function, in which any number of domains may be present. As we shall see below, the answer to such a question is by no means obvious or trivial. This kind of investigation is physically relevant only below the bulk critical temperature, because the probability of two "reversed" domains being near one another is then small. Above the bulk critical temperature, a multitude of domains of all sizes are simultaneously present, and the growth of one is bound to be influenced by the proximity or lack of proximity of others.

4. STATEMENT OF THE PROBLEM, AND STEPS TOWARDS ITS SOLUTION

The actual problem whose solution is required is the enumeration of domains of various perimeters and areas, any boundary of which

(A) does not cross itself;

(B) does not "double back," i.e., an R step may not be immediately followed or preceded by an L step;

(C) does not use any "leg" twice (in the manner shown in Fig. 2). As stated previously, we have obtained what are believed to be correct asymptotic results (valid therefore for very long boundaries) for the case where the perimeters, (but not the areas) of the domains are specified. We shall first consider some similar problems which are distinctly easier to solve, and may be regarded as approximations to the above problem. We shall divide them into groups according to the following criterion.

Group I: Not all of restrictions A, B, and C are properly satisfied.

Group II: Restrictions A, B, and C are all satisfied, and some other constraint is imposed as well.

In Group II we shall clearly get an underestimate of the number of possible paths of given lengths, while Group I will usually give us an overestimate. There is, at present, no known way of telling whether such errors are likely to alter the analytic nature of the singularity of the generating function, to cause a significant shift in the *location* of the singularity, or both. Experience with the Ising model has been that nearly all approximation methods fail badly in *both* these respects. It is hoped that a further study of these rather simpler problems may throw some light on this important question of

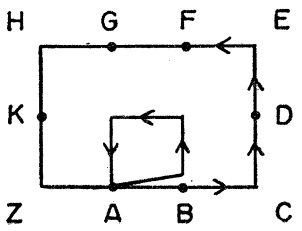

FIG. 2. One type of "forbidden" domain.

finding new criteria for assessing theories of phase transitions from this point of view:

We now consider various simplified versions of the actual problem.

Model O. (Onsager Type Boundary)

We modify function (1) in the following ways: (i) No L steps are ever allowed. (ii) Each group of U or D steps is preceded and followed by at least one R step. As is shown by Temperley,[2] the generating function for a single R step, followed either by further R steps, or by any number of successive U steps, or any number of successive D steps, is

$$x + \frac{xy}{1-y} + \frac{xy}{1-y} = \frac{x(1+y)}{1-y},$$

(R step; R step and U steps; R step and D steps)

so that the generating function for all possible numbers of R steps is

$$\left[1-\frac{x(1+y)}{1-y}\right]^{-1}. \tag{2}$$

(We have dropped the variables δ and ϵ because this boundary cannot enclose a finite domain). This function has a singularity for $x=(1-y)/(1+y)$, which is precisely the same as Onsager's relation[12] for the Curie temperature in the two-dimensional Ising model. x is the Boltzmann factor associated with an unlike horizontal nearest-neighbor pair, the energy of a like pair being taken as zero. In Onsager's[12] notation, $x=\exp(-2H')$. y is the similar factor associated with the vertical interaction and is equal to $\exp(-2H)$.

We are assuming both horizontal and vertical interactions to be in the "ferromagnetic" direction (favoring like pairs of nearest neighbors), so that x and y are both less than unity. For antiferromagnetic interactions, unlike pairs of nearest neighbors are favored, and the type of domain that forms below T_c will have an alternating structure. Domain boundaries will now represent *breaks* in this alternating structure, that is, will contain pairs of *like* nearest neighbors, and the Boltzmann factor associated with a domain boundary is again less than unity. What we are always studying is the question whether the number of ways of realizing a boundary of given length is enough to compensate for the fact that, whichever the sign of the interaction, we always have to expend energy in order to create a domain boundary. If the free energy associated with a boundary vanishes, it is to be expected that a large number of such domains will appear.

Onsager's type of boundary clearly belongs to group II. All three relations A, B, and C are satisfied, but the complete avoidance of L steps is too restrictive. Comparison with the exact problem (model R below) shows that the location of the singularity is correct, but that its analytic form is wrong.

[12] L. Onsager, Phys. Rev. **65**, 117 (1944).

Model P

We prohibit "doubling back," that is, an R step may never be followed by an L step, etc. This means that the path may turn through angles of 0° or 90° at each point, but never the full 180°. We are thus taking *some* account of steric effects between neighbors in a polymer chain but are still allowing portions of the path to cross or to lie side by side. The conditions A and C are therefore not satisfied and the model belongs to group I.

Consider the following 4×4 matrix:

$$P=\begin{array}{c|cccc} & R & U & L & D \\ R & x\delta & \alpha x^{\frac{1}{2}}y^{\frac{1}{2}}\delta & 0 & \alpha^{-1}x^{\frac{1}{2}}y^{\frac{1}{2}}\delta \\ U & \alpha^{-1}x^{\frac{1}{2}}y^{\frac{1}{2}}\epsilon & y\epsilon & \alpha x^{\frac{1}{2}}y^{\frac{1}{2}}\epsilon & 0 \\ L & 0 & \alpha^{-1}x^{\frac{1}{2}}y^{\frac{1}{2}}\delta^{-1} & x\delta^{-1} & \alpha x^{\frac{1}{2}}y^{\frac{1}{2}}\delta^{-1} \\ D & \alpha x^{\frac{1}{2}}y^{\frac{1}{2}}\epsilon^{-1} & 0 & \alpha^{-1}x^{\frac{1}{2}}y^{\frac{1}{2}}\epsilon^{-1} & x\delta \end{array}. \tag{3}$$

In this, we associate the element P_{RR} with an R step that is about to be followed by another R step, the element P_{LU} with an L step about to be followed by a U step and so on. If we write down the nth power of this matrix, the rule of matrix multiplication ensures that each term in the trace of this matrix corresponds to a possible choice of path n steps long. If, for example, an element with second suffix R appears in any term, it can be followed only by one of the three elements P_{RR}, P_{RU}, or P_{RD} with a first suffix R and reference to (3) shows that any such product contains the factor $x\delta$ associated with an R step. If, further, we restrict ourselves to the trace of any power of P, we only allow terms such as $P_{RR}P_{RU}P_{UR}$ whose initial and final suffixes are identical. This term contains the factor $x\delta x^{\frac{1}{2}}y^{\frac{1}{2}}\delta x^{\frac{1}{2}}y^{\frac{1}{2}}\epsilon$ which is the same as the factor $x^2y\delta^2\epsilon$ called for, according to (1), by the three steps RUR. It can readily be verified that, with the choice of terms given by (3), a similar result holds for paths of any number of steps. The factors α and α^{-1} enable us, if we wish, to keep a record of the net number of right-angled bends (those taken in counterclockwise sense being positive) as we travel along the path. For the present, we shall put $\alpha=1$.

We have still to determine how to treat the nondiagonal elements of the matrix P^n. A nondiagonal element does not correspond properly with any path n steps long; for example, consider the element $P^3{}_{RU}$ when $n=3$, which contains, among others, the term $P_{RU}P_{UR}P_{RU}$. This cannot properly be associated with the path OAB shown in Fig. 3(a) because the first and last steps have only been assigned factors $x^{\frac{1}{2}}$, $y^{\frac{1}{2}}$ and also because the factor $\delta\epsilon\delta$ is appropriate to a path beginning at O and ending at A. The diagonal elements *do*, however, all correspond with such paths; for example the term $P_{RU}P_{UR}P_{RR}$ *does* correspond to the path between O and A, because the redundant factor $x^{\frac{1}{2}}$ associated with P_{RR} exactly replaces the $x^{\frac{1}{2}}$ lacked by the leg starting from O. We therefore take the trace of P^n.

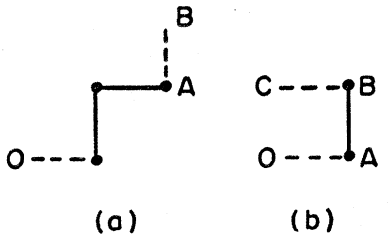

FIG. 3. (a) Typical path enumerated by Tr P^n. (b) Typical path not so enumerated.

This, however, fails to enumerate one type of path, namely that, e.g., beginning with an R step and ending with an L step. In Fig. 3(b) the path $OABC$ calls for factors $P_{RU}P_{UL}$, but we cannot follow these with P_{LR} because no such term exists in the matrix. We can, however take terms such as $(P^{n-1})_{RL}$, multiply each by an appropriate factor ($x\delta^{-1}$ for the element RL), and use these to enumerate paths n steps long, whose initial and final legs are in opposed senses. The complete generating function for all possible lengths of path is thus

$$P(x,y)=1+\mathrm{Tr}\sum_1^\infty P^n+x\delta^{-1}(\sum_1^\infty P^n)_{RL}+x\delta(\sum_1^\infty P^n)_{LR}$$
$$+y\epsilon^{-1}(\sum_1^\infty P^n)_{UD}+y\epsilon(\sum_1^\infty P_n)_{DU}.$$

As long as the generating function converges, we may replace $\sum_1^\infty P^n$ by $P(1-P)^{-1}$, which can be obtained by a straightforward process of inverting the matrix $1-P$. The final results are

$$1+\mathrm{Tr}\sum_1^\infty P^n=\frac{1-x^2-y^2+9x^2y^2-2xy^2(\delta+\delta^{-1})-2x^2y(\epsilon+\epsilon^{-1})}{1+x^2+y^2-3x^2y^2-x(1-y^2)(\delta+\delta^{-1})-y(1-x^2)(\epsilon+\epsilon^{-1})}, \tag{4}$$

$$\text{Remaining terms}=\frac{2xy^2(\delta+\delta^{-1})+2x^2y(\epsilon+\epsilon^{-1})-8x^2y^2}{1+x^2+y^2-3x^2y^2-x(1-y^2)(\delta+\delta^{-1})-y(1-x^2)(\epsilon+\epsilon^{-1})}. \tag{5}$$

Adding these two results together gives us the "path"-generating function. The "domain"-generating function $P(x,y)$ is obtained by applying the operation D (selection of the terms independent of δ and ϵ) to $1+\mathrm{Tr}\sum_1^\infty P^n$. (Incorporating the remaining terms here would mean admitting domains, e.g., that begin with an R step and end with an L step, which are not allowed.) If $x=y$, it is readily seen that the generating function becomes singular as soon as x reaches the value $\frac{1}{3}$. This function does not seem to have any direct application to the ferromagnetic, adsorption, or liquefaction problems, because "domains" in which a given "leg" is used more than once have no direct physical interpretation. It may, however, be applicable to some "long-chain" problems.

Model P'

We might, following the suggestion of Kac and Ward,[4] put $\alpha^4=-1$ in matrix P, and then proceed as before. The object of this is to secure that domain boundaries that cross themselves at right angles are eliminated. This occurs because such boundaries can be described in several senses, which, with the above choice of α, cancel out in pairs. Thus, in Fig. 4, the left-hand method of describing the polygons is associated with the factor unity, while the right-hand one is associated with the factor α^4, so the two cancel out in the generating function. For a similar reason, unwanted domains of the type shown in Fig. 5 are eliminated, but it is readily seen that domains of the type shown in Fig. 2 are *not* eliminated. The inner domain in Fig. 2 is necessarily always described in the same sense as is the outer domain, because, if they were described in opposite senses, we could not avoid having an R step immediately followed by an L step when we came to the particular leg AB in Fig. 2 that is used twice. Furthermore, with this form of generating function, the presence of the α's introduces a factor -1 for each rotation though 2π and the number of boundaries of the type shown in Fig. 2 (enclosing an even number of separate domains) is *subtracted from*, not added to, the number of paths in which the number of separate domains (each touching one or more others along at least one leg) is odd. Thus, the freedom from "cross-overs," and from boundaries such as that shown in Fig. 5 has been somewhat dearly bought, and no physical application of this generating function seems possible unless we can find a way of eliminating diagrams such as that shown in Fig. 2. It is believed that this can be done, at least for the limiting case of a large domain, but we defer description of the method used until we have discussed another model. Meanwhile, we

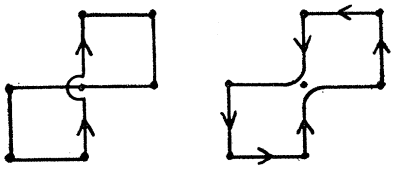

FIG. 4. Mutual cancellation of certain types of "forbidden" domain.

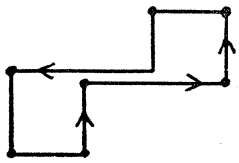

FIG. 5. Another type of "forbidden" domain.

record the result for closed domains described by P'. (We have not considered the use of this function for enumerating *paths*, because paths that *spiral* round the origin would clearly not be enumerated properly because of the presence of the factors α and α^{-1} which introduce a factor -1 each time the path rotates through 2π.)

$$P'(x,y)=D\left(\frac{1-x^2-y^2-3x^2y^2-2xy^2(\delta+\delta^{-1})-2x^2y(\epsilon+\epsilon^{-1})}{1+x^2+y^2+x^2y^2-x(1-y^2)(\delta+\delta^{-1})-y(1-x^2)(\epsilon+\epsilon^{-1})}\right). \tag{6}$$

This function still has a singularity, but it is of a less violent type than those previously encountered. The denominator now vanishes *only* when $\epsilon=\delta=1$ *and* $y=(1-x)/(1+x)$, that is, it only vanishes at the single point $\omega_1=\omega_2=0$ in the range of integration prescribed by the operation D. The numerator in (6) vanishes at this point also. The effect of this is that the generating function is itself continuous everywhere $(0<x, y<1)$ and that its first derivatives with respect to x or y are also continuous, but become logarithmically infinite when $y=(1-x)/(1+x)$. The analytic behavior of this domain-generating function is thus, as one might expect, closely similar to that of Onsager's[12] partition function for the Ising model. The "mean boundary length" of a domain thus becomes infinite at the critical temperature since it depends on the first derivatives, with respect to x and y, of the logarithm of the generating function.

If we regard (6) as an approximation to the true domain-generating function, it is clear that it satisfies restrictions A and B alone but not restriction C, in that domains of the type shown in Fig. 2, containing any number of "loops," are allowed. Nevertheless, it seems certain that expression (6) underestimates the entropy, that is, underestimates the number of single domains of a prescribed path-length. In expression (6), domains consisting of a single loop are reckoned negatively (these are the ones sought) while those consisting of two loops are reckoned positively, those of three loops negatively and so on. For a given length of perimeter the number of domains of two loops almost certainly exceeds the number with three loops and so on. If this is so, the numerical value of any coefficient in expression (6) (the sum over loops of all sizes) is less than the number of single-loop domains.

Before considering the question of the removal of these unwanted "multiloop" domains from the generating function, we shall consider another model, which is admittedly an approximation to the true domain generating function that we are looking for, but, seems to "mimic" its behavior extremely closely, not only reproducing many of the early terms accurately, but apparently having a singularity in the correct place and of nearly the correct analytical type. The smallest two-loop domain consists of a single square described twice, which means that expression (6) reproduces the terms of the exact generating function only as far as terms like x^2y^4 but gives the wrong coefficient for x^4y^4, while the model about to be considered does not break down until we reach a term like x^6y^6.

Model Q

We can obtain definite results for a particular type of domain in the square lattice, by a somewhat different method, as follows: We restrict ourselves to domains which can be built up column by column as follows: (a) Each column consists either of a single square, or of a number of squares in a continuous straight line. (b) Each column must overlap those to the left and right of it by at least one lattice distance. These conditions clearly secure that we have a *connected* domain, but are too restrictive, because it is obviously possible to have a connected domain with *gaps* in some or all of the columns. Nevertheless, this model is of interest, because a record can be kept of both the *area* and *perimeter* of the domain, and many results can be obtained in finite terms. The model clearly belongs to group II, as it is too restrictive.

Let g_r be the generating function for a domain whose left-hand column contains exactly r squares. The next column to the right may contain 0, 1, 2, $\cdots$ squares. The left-hand column, by itself, would call for a term $x^2y^{2r}z^r$, because its perimeter contains two horizontal lines and $2r$ vertical ones. If, however, the next column to the right is a single square, it can touch the first one in r different places, and the two together would call for a term $rx^4y^{2r}z^{r+1}$. If, however, this column were a doublet and $r>1$, the two columns could overlap in $r+1$ different ways. In two of these cases, the perimeter is described by the factor x^4y^{2r+2} while in the remaining $r-1$ cases, the perimeter is described by x^4y^{2r}. Proceeding in this way, imagining the domain built up, column by column from left to right, and taking account of the various possible overlaps between adjacent columns, we find the following set of equations for the g_r's.

$$\begin{aligned}
g_1&=x^2y^2z+x^2z[g_1+2g_2+3g_3+4g_4+\cdots],\\
g_2&=x^2y^4z^2+x^2z^2[2y^2g_1+(1+2y^2)g_2+(2+2y^2)g_3+(3+2y^2)g_4+\cdots],\\
g_3&=x^2y^6z^3+x^2z^3[3y^4g_1+(2y^2+2y^4)g_2+(1+2y^2+2y^4)g_3+(2+2y^2+2y^4)g_4+\cdots],\\
g_4&=x^2y^8z^4+x^2z^4[4y^6g_1+(3y^4+2y^6)g_2+(2y^2+2y^4+2y^6)g_3+(1+2y^2+2y^4+2y^6)g_4+\cdots].
\end{aligned} \tag{7}$$

It is readily seen that these quantities satisfy the following difference equation arrived at by subtracting the equation for zy^2g_r from that for g_{r+1}, and repeating the process on the resulting equations:

$$g_r-2y^2zg_{r-1}+y^4z^2g_{r-2} = x^2z^r(1-2y^2+y^4)(g_r+2g_{r+1}+3g_{r+2}+\cdots). \quad (8)$$

Applying a similar process twice more, the factor this time being z, we get:

$$g_{r+2}-2z(1+y^2)g_{r+1}+z^2(1+4y^2+y^4)g_r -2z^3(y^2+y^4)g_{r-1}+z^4y^4g_{r-2}=x^2z^{r+2}(1-y^2)^2g_r. \quad (9)$$

These equations are completely soluble in finite terms in two important special cases: $y=1$, corresponding to the case where we do not specify the total perimeter but the x and z variables remain available to specify the number of columns and the total area of the domain, and $z=1$, in which we no longer specify the area of the domain, but *do* specify its perimeter.

For $y=1$, we have, from Eq. (8),

$$g_r-2zg_{r-1}+z^2g_{r-2}=0,$$

the general solution of which is

$$g_r=A(z,x)z^r+B(z,x)rz^r \quad (10)$$

by an argument similar to that used for linear differential equations. A and B are arbitrary functions of x and z but are independent of r. They can be determined from the first two of Eqs. (7) as follows: Define G and H as follows, $G=\sum_r g_r$, $H=\sum_r rg_r$ (G is the complete generating function for a domain with a first column of any length. g_r is the generating function if the first column is *known* to be r units long). From Eq. (10) we have

$$g_1=zA+zB, \qquad G=Az/(1-z)+Bz/(1-z)^2,$$
$$g_2=z^2A+2z^2B, \quad H=Az/(1-z)^2+Bz(1+z)/(1-z)^3,$$

From Eqs. (7) (first two), we get

$$g_1=x^2z+x^2zH,$$
$$g_2=x^2z^2+x^2z^2(G+H),$$

leading to the following expression for G:

$$G=\frac{x^2z(1-z)^3}{1-(x^2+4)z+(x^2+6)z^2-(x^4-x^2+4)z^3+(1-x^2)z^4}. \quad (11)$$

It is remarkable that the *exact* generating function for this fairly complicated problem can be expressed in closed algebraic form. If we put $x=1$, this expression diverges when z is approximately $\frac{1}{4}$. This result seems to give us insight into the reason for the curious result found by Kirkwood and Monroe.[13] According to this result, a liquid consisting of rigid spheres may be expected to "freeze" into an ordered lattice as soon as its density exceeds a certain value, even though the spheres do not attract one another. We attempt to represent the liquid by the lattice model (Yang and Lee),[8] by dividing our vessel up into cells, and distinguishing only between configurations in which a given cell is empty or contains just one molecule. Result (11) appears to mean that, if the probability of occupation of each cell exceeds a certain value, there will be a gain in entropy if two or more isolated domains join up to form one connected domain. It also seems to be closely related to the fact that a branching polymer may be expected to form a gel if the probability of forming branches exceeds a certain value.[14]

We may use Eq. (9) in the special case in which $z=1$, that is, in which we are enumerating domain boundaries but not domain areas. As before, we look for solutions of the type $g_r=A\lambda^r$, and obtain the following equation for λ:

$$(\lambda-1)^2(\lambda-y^2)^2-\lambda^2x^2(1-y^2)^2=0. \quad (12)$$

It is readily shown that λ never attains unity for any relevant values of x and y, and that any pair of complex conjugate roots necessarily has a modulus equal to y which is less than unity. Now the most general possible solution of (9) leads to the result

$$G(x,y)=\sum_r g_r=A_1\lambda_1/(1-\lambda_1)+A_2\lambda_2/(1-\lambda_2)+\cdots,$$

where $\lambda_1\cdots\lambda_4$ are the roots of Eq. (12) and A_1, A_2, A_3, A_4, are arbitrary functions, which can be determined from the first four of Eqs. (7). Now, in view of the above information about the λ's, it might appear at first sight that the generating function $G(x,y)$ can only diverge, if at all, for a reason analogous to that leading to the divergence of G in (11), that is as a result of the form of functions $A_1\cdots A_4$. However, the function $G(x,y)$ *can* have an infinite derivative with respect to x or y if two roots of Eq. (12) become equal, the condition for which is

$$1+y^2\pm x(1-y^2)=\pm 2y,$$

$x=(1-y)/(1+y)$ being the only case that is physically relevant. If x is near this critical value $x_c-\delta$ (say), it is easily shown that, although the λ's are all finite, $\partial\lambda/\partial x$ is proportional to $\delta^{-\frac{1}{2}}$ for the two λ's that are nearly equal. Thus, G itself is continuous, but it has infinite derivatives as x or y passes through the critical value.

It is most unlikely that any singularities have been missed, because this model is, in fact, based on taking just the domains that can be enclosed between *two* Onsager-type boundaries (model O) which are permitted to cross at no more than two points and the only

[13] J. G. Kirkwood and J. E. Monroe, J. Chem. Phys. **9**, 514 (1941).

[14] P. J. Flory, *Principles of Polymer Chemistry* (Cornell University Press, Ithaca, 1953), Chap. 9.

singularities for model O are known. The model seems worth considering, because of the fact that so many results can be obtained in finite terms, and also because its generating function is probably a close "mimic" of the true one. Not only does it, as we shall see below, diverge at the same place as the true one, but it reproduces a great many of the early terms of the true generating function exactly. The smallest domain that is not enumerated by this model is shown in Fig. 6, which calls for a factor $z^5x^6y^6$; all previous domains are reproduced accurately. In this respect it is decidedly better than model P', which also diverges at the right place but already departs from the true generating function at the term $z^2x^4y^4$ (a single square being spuriously encircled twice, according to this latter model).

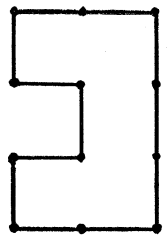

FIG. 6. Smallest domain not enumerated by Model Q.

Models R and S. The Exact Problem

We now show that model P' can be improved further, confining ourselves to the case $y=x$. The generating function (6) for domains is incorrect for two reasons: (a) It permits the same domain to be described any number of times in the same sense (each fresh repetition introducing a new factor -1 because of the presence of the α's). Thus, a single square, corresponding to a term x^4 in the true generating function, contributes to function (6) the terms $-x^4+x^8-x^{12}+x^{16}-\cdots$. (b) It allows "multiloop" domains of which Fig. 2 is a typical example.

We now assume that we possessed the exact generating function for "single loop" domains, and then ask how this function would have to be modified in order to reproduce the *known* generating function for model P', which we call $g(x)$ for the case $x=y$ and which is defined by Eq. (6). In other words we shall try to derive a *functional equation* for the unknown generating function in terms of the known one. It is convenient to introduce a slight generalization of our problem at this stage. We define the true generating function $f(u,v)=\sum_m \sum_n A_{mn}u^m v^n$, where A_{mn} is the number of single-loop domains whose perimeter has n right-angled bends (corresponding to elements like RU, DL, etc., in matrix P or P') and m "straights" (corresponding to elements like RR, LL, etc.) We confine our attention to just one such domain, N steps long, which we call domain N, which corresponds to just one term $(m+n=N)$ of $f(u,v)$, and set up a correspondence between this term and some of the terms of $g(x)$. To allow for the possible repeated description of domain N, together with the alternation of signs introduced by $g(x)$, we replace $f(u,v)$ by an associated function $h(u,v)$ defined by $h(u,v)=\sum_m \sum_n[A_{mn}u^m v^n/1+u^m v^n)]$. We call this intermediate model, in which "single loop" domains may be described repeatedly, but domains such as those shown in Fig. 2 are still not allowed, as model S. The two generating functions are related in the same way as are Taylor and Lambert series, and, if one series is known for all m and n up to some given value of $m+n$, all the corresponding terms of the other can be calculated.

We now allow for the fact that $g(x)$ allows each of the points of domain N to be made the starting point from which further closed domains can be described in the manner shown in Fig. 2. Thus, if the outer rectangle of Fig. 2, with Z as the starting and finishing point, is domain N, we may, according to $g(x)$, *leave* domain N at a point such as B, then describe an additional closed domain such as the small square, finally returning to B in such a way that the *last* leg, AB, of the additional domain belongs to domain N. After returning to B, we continue to describe domain N. When we arrive at C we have a similar option of proceeding to D, or of leaving N, describing another closed domain, returning to C and then proceeding to D, and so on. In Fig. 2, we consider that the subsidiary domain "begins" at B rather than at A, because, when we arrive at A we have to proceed to B whether we are going to describe the small square or remain on N, and it is not until we arrive at B that we have to leave N and begin to describe the small square.

It is already clear that any one term in $f(u,v)$ is going to be associated with infinitely many terms in $g(x)$, because any point in the typical domain N can be made the starting and finishing point of a subsidiary domain, and the same is true of any point in Fig. 2, and so on for more and more complicated domains. We shall show that all the possible subsidiary domains that begin and end at a point such as B in Fig. 2, can be described by an appropriate selection of terms from $g(x)$ [not the whole of $g(x)$ because there are restrictions on the directions of the starting and finishing "legs"], and we shall show how to write down these subsidiary generating functions in terms of the matrix P'. This process has to be carried out for each point in the typical domain N and the result summed over all such domains in order to arrive at $g(x)$.

At this point we should remark that boundaries which cross themselves make the same total contributions (if any) to both $g(x)$ as originally defined, and to $g(x)$ as we are proposing to construct it from $f(u,v)$. To begin with, it is clear that graphs such as that shown in Fig. 2 are included in both cases, and these such as Figs. 4 and 5 are excluded. Now consider Fig. 7 (the description of the figure being supposed to begin at the bottom left hand corner). If the common leg is described twice in *opposite* senses, there is *no* total contribution to $g(x)$ if we consider all possible methods of drawing this graph

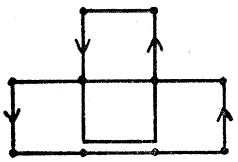

FIG. 7. "Forbidden" domain enumerated by $g(x)$.

(just as there is none from Fig. 5), nor is there any term in $f(u,v)$ corresponding to a primary domain from which we could construct any of these graphs according to the rules laid down. If the common leg is described twice in the *same* sense, then the other legs must be described in the senses shown by the arrows in Fig. 7 (or all in exactly the opposite senses). Now Fig. 7 can be split up into a "primary" and a "subsidiary" domain in four different ways. These contribute four different terms to $g(x)$ because they correspond to different permutations of the same set of legs, and arise from four different terms of $f(u,v)$, because each of them represents a different choice of primary domain. Again, graphs such as Fig. 8 are retained by both constructions of the generating function, subject to the proviso that the smaller domains are described *separately*, e.g., we start at the bottom left-hand corner, proceed to A, describe one of the smaller domains, return to A, proceed to B along the outer domain, describe the other smaller domain, return to B, and then complete the outer domain. It can be verified that some paths, which correspond to the *unicursal* description of the two smaller domains (e.g., starting at A and returning to A after describing both), cancel out in pairs, while others persist in both forms of the generating function.

While it has not yet been formally proved, the (1–1) correspondence between the terms of $g(x)$ and the proposed equivalent generating function has been verified in a large class of cases. It seems unlikely that the exceptional cases, even if any exist, will be numerous enough to upset the asymptotic results. Our work suggests that $g(x)$ might be obtainable from $f(u,v)$ or $h(u,v)$ by a transformation of the following kind:

$$u,\ v=x\{1+\tfrac{1}{4}g(x)\}, \tag{13}$$

which we arrive at by the following argument. The outer domain in Fig. 2 represents a typical term of $f(u,v)$. Suppose that we have reached B in Fig. 2. The first term, x, in the proposed expression (13) for u or v, corresponds to proceeding just to C in Fig. 2. A term $xg(x)$, inserted into f in place of u, would correspond to describing any closed domain starting and finishing at B, then proceeding to C. The factor $\frac{1}{4}$ allows for the fact that not all possible closed domains enumerated by $g(x)$ are suitable for introduction between the legs AB and BC, since we are limited to domains whose initial leg differs from BC but can follow AB, while the final leg must be capable of being followed immediately by BC. In Fig. 2, the initial leg must be U or D and the final leg must not be L nor must it be opposed to the initial leg. Even now, the argument is not quite right, and we have to improve it further. In the first place, care is necessary with the terms in $h(u,v)$ corresponding to repetition of the same domain. It would not be correct to apply the transformation (13) to u and v wherever they appear in $h(u,v)$. Consider, for example, a term in $g(x)$ corresponding to p repetitions of the same domain N, a departure from this domain during the $(p+1)$th circuit, the tracing of a subsidiary domain, the completion of the $(p+1)$th circuit of N, followed by q further circuits of N. If domain N contains m "straights" and n "angles," it is described by one term, $u^m v^n$, from $f(u,v)$. If a subsidiary domain is put in during the first circuit, we apply a transformation such as (13) to one of the factors u or v. If, however, the domain N is described p times before we first leave it to form a subsidiary domain, the corresponding term in $g(x)$ is $(x^m x^n)^p u^m v^n$. If we now apply transformation (13) to u and v, this takes account of the fact that subsidiary domains may be formed anywhere during the $(p+1)$th circuit of N. It would, however, be redundant to take explicit account of any of the q *subsequent* circuits of N. For the term $\frac{1}{4}xg(x)$ in (13) automatically describes, not only the introduction at any point of N of a single subsidiary domain, but also the introduction of *any* closed domain, including, as a particular case, a single subsidiary domain followed by q further circuits of N. A little consideration shows that all possibilities are covered if we write,

$$g(x)=\sum_m \sum_n \frac{A_{mn}u^m v^n}{1+x^{m+n}}, \tag{14}$$

where we have still to determine the correct transformations of u and v to describe the insertion of subsidiary domains. They are of the same type as expression (13), but slightly more complicated. Expression (14) implies that, in the event of *repeated* description of a domain N, it is only the complete circuits that occur *before* the first departure from N that need be specified explicitly, the remaining circuits, if any, already being implied by the introduction of $g(x)$ into (13).

There is still one remaining difficulty that forces us to modify expression (13), and, in fact, we find that the required transformation is slightly different for the "straights" and "angles" in the typical domain N. To

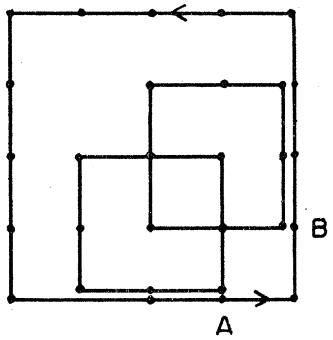

FIG. 8. "Forbidden" domain giving a finite contribution to $g(x)$.

see the nature of the difficulty, consider a graph such as that shown in Fig. 9, in which a subsidiary domain has two or more *successive* legs in common with domain N. (A graph such as that shown in Fig. 10 does not cause this trouble, and is covered by what we have already done.) The difficulty is that, in Fig. 9, the extra loop could, according to the transformation (13), be inserted between legs BC and CD *or* between CD and DE of the domain N, and these two possibilities would be reckoned separately if we applied (13) to expression (14), though both of them correspond to precisely the same set of successive choices of terms from the enumerating matrix P'. Consider the terms of P' corresponding to passing from A to E in Fig. 9, taking in the subsidiary domain en route.

Leg	AB	BC	CD	DH	HI	IB	BC	CD	DE
Term of P'	DR	RR	RR	RU	UL, LL	LD	DR	RR	RR

This choice of terms is independent of whether we suppose the subsidiary domain to "start" at C or at D, whereas applying (13) to (14) would list these two cases as separate possibilities. We therefore modify (13) so that the starting point of the subsidiary domain is uniquely located at D in Fig. 9 but so that any point in domain N can still be used as a starting point for a new domain.

The required correction can be made as follows. We are deriving all the terms of $g(x)$ corresponding to a particular "main" domain N. In deriving them, we have, at each step of N, the option of remaining on N or of beginning a subsidiary domain. Now, we cannot say that we have "begun" a new domain until we have "departed" from the old one, that is, until we have reached a point such as D in Fig. 9 and taken a step like DH which is definitely different from DE. Thus, given the steps CD and DE in the large domain N, we wish to enumerate all those domains for which:

(a) The first step is different from DE. In Fig. 9 it has to be a U or D step. (It cannot be an L step because this cannot follow CD.)

(b) The last step can close on to the first step to make a domain, i.e., if the first step is U, the last step must not be D. Also, it must be followed directly by DE. In Fig. 9 this means that we want all the domains, specified by P' or $g(x)$ which can be followed by an R step and begin with either a U or a D step. We can select these in the same way that we selected domains of various kinds enumerated by the matrix P' or P. (The discussion is less complicated, because we are not now

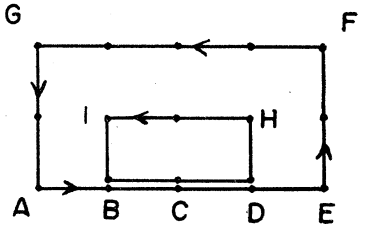

FIG. 9. Subsidiary domain has successive legs in common with domain N.

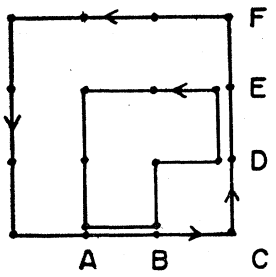

FIG. 10. Subsidiary domain has nonsuccessive legs in common with domain N.

interested in *open* paths, and we do not encounter the difficulties associated with the end steps not having their proper factors.) The possible subsidiary domains must begin with P'_{RU} or P'_{RD} and must be closed and must not end with an L step. Thus, for the point D in Fig. 9 (a "straight") the R step and possible subsidiary domains are enumerated by

$$P'_{RU}D[(1-P')^{-1}]_{UR}+P'_{RD}D[(1-P')^{-1}]_{DR}, \quad (15)$$

and, for such a point, we can replace $\frac{1}{4}g(x)$ in (13) by the above expression. P'_{RU} specifies an R step about to be followed by an upward step, while $(1-P')_{UR}^{-1}$ specifies paths of arbitrary length beginning with a U step, the end step about to be followed by an R step, while application of the operator D restricts us to closed domains.

We can carry through a similar discussion for the various possible types of point that may occur in a domain boundary, corresponding to the twelve non-vanishing elements of P'. In the case $x=y$, expression (15) is numerically the same for all four types of "straight" points (described by the elements RR, LL, etc., of P'), and we find that a similar result holds for all the "angle" type points. Consider, for example, an "angle" of the type RU. If a subsidiary domain is to begin at such a point, its first leg must be either R or D (i.e., different from U but not L, since L cannot follow immediately after R), while the last leg must be R. For this case (15) is replaced by

$$P'_{RR}D[(1-P')^{-1}]_{RU}+P'_{RD}D[(1-P')^{-1}]_{DU}, \quad (15)'$$

which is to replace $\frac{1}{4}xg(x)$ in (13).

Evaluating the elements of $(1-P')^{-1}$, and using expressions (15) and (15)′ to correct transformation (13), we find

$$\begin{aligned} u&=x\Big[1+\frac{1}{4\pi^2}\int_0^{2\pi}\int_0^{2\pi} \times\frac{d\omega_1 d\omega_2(2x^2+2x^2\cos\omega_1\cos\omega_2)}{(1+x^2)^2-2x(1-x^2)(\cos\omega_1+\cos\omega_2)}\Big], \\ v&=x\Big[1+\frac{1}{4\pi^2}\int_0^{2\pi}\int_0^{2\pi} \times\frac{d\omega_1 d\omega_2(x^2+x^2\cos\omega_1\cos\omega_2)}{(1+x^2)^2-2x(1-x^2)(\cos\omega_1+\cos\omega_2)}\Big], \end{aligned} \quad (16)$$

TABLE I. Asymptotic behavior of domain-generating functions.

Model	Generating function	x_c	Asymptotic behavior of generating function ($x=x_c+\delta$)	Remarks
Unrestricted	Expression (1)	$\frac{1}{4}$	$\log\delta$	
P	Expressions (4) and (5)	$\frac{1}{3}$	$\log\delta$	"Doubling-back" forbidden
Q	$G(x,x)$. Expression (9)	$\sqrt{2}-1$	$\delta^{+\frac{1}{2}}$	Some domains omitted
P'	$g(x)$. Expression (6)	$\sqrt{2}-1$	Constant $+\delta\log\delta$	"Multi-loop" domains
S	$h(x,x)$	$\leqslant\sqrt{2}-1$	Uncertain	Repetitions allowed
R	$f(x,x)$	$\leqslant\sqrt{2}-1$	Uncertain	

the desired functional equation for $f(u,v)$ then being found by substitution of these relations into (14) leading to $g(x)=f(u,v)-f(ux,vx)+f(ux^2,vx^2)-\cdots$. The resulting relation seems to be the only piece of analytic information that is at present available about this generating function, f. It is not theoretically sufficient to determine it, because we have had to introduce the two variables u and v into the definition of f, whereas $g(x)$ only involves one variable. We do not know the function corresponding to $g(x)$ in the case where "straights" and "angles" occur with differing probabilities because the mutual cancellation of domains like those shown in Figs. 4 and 5 only holds strictly when these probabilities are equal. The more general generating function is, however, probably closely related to the corresponding generalization of the matrix P', which we put on record in the Appendix on account of possible physical applications.

5. ANALYTIC BEHAVIOR OF THESE GENERATING FUNCTIONS NEAR THEIR SINGULARITIES. COMPARISON WITH NUMERICAL EVIDENCE FROM MACHINE RUNS

It is of interest to compare the analytic behavior of these domain generating functions as we apply various types of constraint. The evaluation of the double integrals (corresponding to the operation D) in terms of elliptic integrals is standard and calls for no comment. (The singularities arise from the logarithmic divergence of the integral $K(k)$ as $k\rightarrow 1$.) We exhibit the results in Table I, omitting numerical coefficients and confining ourselves to the case $x=y$. In Table I, x_c is the critical value of the selector variable x. These results make it extremely probable that the singularity is in the same place for the domain-generating function and for the corresponding Ising problem. Further light is thrown on the corresponding question for other lattices by the work of Wall and others[9,10] on the generation of non-intersecting random walks on various lattices. It is shown in Sec. 8 that the number of such domains, or paths, of N legs is asymptotically proportional to x_c^{-N}, where x_c is the singularity in the domain generating function. $1/x_c$ may thus be regarded as the "effective number of choices" associated with the particular lattice. At every step, a certain actual number of choices is possible (3 in the plane square lattice if the backward step is excluded) but a certain percentage of all the paths then intersect themselves and have to be deleted. The above work indicates that this effective number of choices is almost certainly $1/(\sqrt{2}-1)=2.412\cdots$. Wall and others[9,10] define an "attrition coefficient" describing this loss of paths at each step, and they find that it settles down to a nearly constant value while the paths are still quite short. It is of interest to compare the x_c's calculated from the attrition coefficients (the attrition coefficient being the ratio of the effective number of choices to the actual number of choices) with the positions of the singularities estimated for the corresponding Ising lattices [from data collected by Domb and Potts.[15]] In their notation, $x_c=\tanh(J/kT_c)$. This comparison is shown in Table II.

The agreement is in all cases reasonable, bearing in mind the fact that the exact value of $1/x_c$ is known only for certain two-dimensional lattices and the values in some of the three-dimensional cases are in considerable doubt. The values from the machine are nearly all higher than these for the Ising model, and this *is* in the direction consistent with the existence of lower critical temperatures of Mayer type. Equally it could mean that the "measured" attrition coefficients are all slightly higher than the limiting values appropriate to very long paths.

The process used by Montroll[16] removes only paths containing one or more simple square loops, leaving all other types of self-crossing paths containing loops of more than four legs each. It therefore only gives an upper bound to the entropy.

6. THE "SPREAD" OF A LONG PATH AS A FUNCTION OF THE NUMBER OF "LEGS"

In the generating function (1) (no constraints), the direction of each step is completely independent of what previous steps may have been taken; generating function (2) describes a path in which one direction of step is forbidden completely, while function (5) introduces correlations between successive steps, "doubling-back" being forbidden. Generating function (6) introduces what one might call "long-distance" correlation. As we have seen, this function does not enumerate any path that crosses itself at right angles, so that the direction of any given step depends, to some extent, on *all* the steps that preceded it. Considerable study has been given to

[15] C. Domb and R. B. Potts, Proc. Roy. Soc. (London) **A210**, 125 (1952).

[16] E. W. Montroll, J. Chem. Phys. **18**, 734 (1950).

the question of how constraints of this last kind may be expected to affect the mean "spread" (distance between the two ends) of paths consisting of a fixed number of legs. It is obvious that, of two paths of equal length, one with small spread is more likely to cross itself than is one with large spread, and therefore that the introduction of such constraints will increase the mean spread.

For the completely unconstrained path, it is well known that the mean spread is proportional to the square root of the number of legs when this number becomes large, and it is of some interest to decide whether the introduction of constraints merely alters the constant of proportionality in the law

$$(\text{spread})^2 \propto \text{total length}, \tag{17}$$

or whether the analytic form of the law is changed. This question arises in the study of high polymers, where self-crossing configurations are not physically permissible. The theoretical and numerical evidence bearing on this question has recently been summarized by Wall *et al.*[9,10] It is probable, but not quite certain, that the situation is essentially different in two dimensions and in three dimensions, as is known to be the case for a completely random path. We study the asymptotic behavior of the spread of long paths called for by the path-generating functions that we have described above.

Since a generating function describes all possible paths, it contains implicitly the desired information about spread, and it seems worth showing explicitly some of the possible ways in which departure from a law of the type (17) can be associated with the analytic form of the generating function. If we put $x=y$, the most general type of generating function that we have met with in this paper may be expressed in the form

$$\frac{\chi(x)+\psi(x)(\cos\omega_1+\cos\omega_2)}{1-\phi(x)(\cos\omega_1+\cos\omega_2)}, \tag{18}$$

where χ, ψ, and ϕ are algebraic functions of various kinds. To illustrate the method of deriving the relation between spread and path length, we refer first to the one-dimensional problem. We may suppose the generating function expanded in the form

$$l(x,\epsilon)=\sum_{n,m} p_{n,m}x^n\epsilon^m, \tag{19}$$

where, as before, n represents the total number of, steps and m the algebraic separation of the ends of the path. By symmetry, we necessarily have $p_{n,m}=p_{n,-m}$. In practice, if we have a generating function as complicated as (18) it may be necessary to resort to the method of steepest descents in order to get asymptotic estimates of $p_{n,m}$, which is the first step in calculating the mean spread. (For the generating functions considered in this paper, the desired results can be obtained algebraically.) Equation (19) is equivalent to the statement

$$\sum_m p_{n,m}\epsilon^m=\text{coeff of } x^n \text{ in } l(x,\epsilon). \tag{20}$$

Differentiating Eq. (20) twice with respect to ϵ, we have

$$\sum_m m^2 p_{n,m}\epsilon^m=\left(\epsilon\frac{\partial}{\partial\epsilon}\right)^2[\text{coeff of } x^n \text{ in } l(x,\epsilon)].$$

What we want to calculate as a function of n is the mean value of m^2, averaged over all possible paths of total length n, that is,

$$\langle m^2\rangle_{\text{Av}}=\sum_m m^2 p_{n,m}/\sum_m p_{n,m} \quad (\text{fixed } n). \tag{21}$$

The denominator and numerator of this fraction can be calculated from expressions (20) and (21), respectively, ϵ being set equal to unity after the differentiations. If $\epsilon=e^{i\omega}$, $(\epsilon\partial/\partial\epsilon)^2\equiv-\partial^2/\partial\omega^2$. This work is readily extended to two or three dimensions, the square of the length of any path simply being the sum of m_x^2, m_y^2 and m_z^2.

Carrying out this work for the generating function given in (18), we find

$$\sum_m p_{n,m}=\text{coeff of } x^n \text{ in } \frac{\chi+2\psi}{1-2\phi},$$

$$\sum_m m^2 p_{n,m}=\text{coeff of } x^n \text{ in } \frac{2\chi\phi+2\psi}{(1-2\phi)^2},$$

and the expansions can be performed by partial fractions if ϕ, χ, and ψ are known algebraic functions.

TABLE II. Comparison of machine and analytic results for various lattices.

Lattice	Attrition coefficient	Estimates of $1/x_c$ By machine	From Ising model	Other
2 choice square	0.798	1.60	...	<1.62 (Montroll) 1.41 (Temperley, unpublished)
3 choice square	0.888	2.66	2.41 (exact)	≥2.41 (this paper) <2.77 (Temperley, Montroll method)
2 choice honeycomb	0.931	1.86	1.73 (exact)	
5 choice triangular	0.840	4.20	3.37 (exact)	
4 choice simple cubic	0.902	3.61	3.08 (interpolated)	
5 choice simple cubic	0.941	4.71	4.71 (estimated from series)	
3 choice diamond	0.972	2.92	2.41 (interpolated)	
7 choice body-centered	0.933	6.53	5.67 (estimated from series)	
11 choice face-centered	0.917	10.09	10.49 (estimated from series)	

TABLE III. Comparison of entropy and spread for various path-generating functions.

Model	$\Sigma_m p_{m,n}$	$\Sigma_m m^2 p_{m,n} x^n$	$\langle m^2\rangle_{Av}$	Entropy
Unrestricted	$1/(1-4x)$	$4x/(1-4x)^2$	n	$k \log 4$
P	$(1+x)/(1-3x)$	$4x(1+x)/(1-x)(1-3x)^2$	$2n$	$k \log 3$
P'	$(1+2x+3x^2)/(1-2x-x^2)$	$\dfrac{4x(1+x^2)(1+2x-x^2)}{(1-2x-x^2)^3}$	$\frac{1}{2}n^2$	$k \log(1+\sqrt{2})$

In Table III we compare three different types of path, the unrestricted path [expression (1)], the path in which doubling back is forbidden [model P, expressions (4) and (5)], and the path in which "crossovers" are prevented [model P', expression (6)]. We have seen that this last expression is not satisfactory as a path-generating function (because certain types of looped path are still included, and also because paths that spiral one or more times around the origin will not be properly counted), but it is of interest to examine expression (6) as if it *were* a proper path-generating function, as we thereby gain insight into possible mathematical peculiarities of the generating function which would imply breakdown of law (17). The different analytic behavior of models P and P' is associated with the fact that, in model P, the smallest positive root of $1-2\phi=0$ is a single root, while in model P' it is a repeated root. This is the same peculiarity that is responsible for the different analytic behavior of the domain-generating functions near the critical temperature that we have already discussed. Quite similar behavior is to be expected for the two-dimensional triangular and honeycomb attices. Wall *et al.*[9,10] give evidence that, in the two-dimensional lattice $\langle m^2\rangle_{Av}/n$ *does* diverge in two-dimensional lattices if we reject "loop" paths.

Our conclusion is the common sense one that, in two or more dimensions, short-range correlation between legs, e.g., between successive legs only [as in generating function (4)], cannot affect the *form* of the law (17)[17] (though it does increase the numerical coefficient), but that an alteration in the form of the law may quite easily occur if we introduce a correlation between a given leg and *all* preceding legs, as in generating function (6), model P'. The latter type of correlation is what is called for by the physics of high-polymer chains.

7. COMPARISON OF VARIOUS APPROXIMATIONS

Inspection of Table I shows that, as the approximation is progressively improved, we first get an improvement in the location of the transition temperature, then the singularity approaches more and more nearly to the right form. Concurrently, we find that the approximate generating function reproduces more and more of the early terms of the exact series, though an approximation that reproduces a great many of the early terms is *not* necessarily the one that best describes the behavior near the transition temperature. (Kramers and Wannier,[18] discussing various approximations to the Ising partition function, also found that they could not be placed in any *unique* "order of merit".) It is of interest to compare these successive approximations with some recent work on the liquid distribution function by Rushbrooke and Scoins,[19] and by Nijboer and Fieschi.[20] Better treatments, reproducing correctly Mayer cluster integrals of progressively higher order, have been developed by these authors, and have been checked by comparison with the "rigid sphere liquid" for which some precise numerical information is available. The present position is somewhat disappointing, as it seems that each successive improvement in the approximate methods used only results in reproducing about one more term of the virial series correctly, so we are still almost without information on the analytic form of the singularity in the virial series, which should correspond to the onset of liquefaction.

8. ENTROPY OF PATHS AND DOMAINS. PROBABILITY OF RING CLOSURE

The entropy of a system such as a single magnetic domain, a connected cluster of adsorbed atoms, or a gas-bubble in a liquid can be calculated according to this type of model, simply by identifying the corresponding generating function with the partition function for a single domain, the selector variable being associated with the appropriate variable, e.g., magnetic field, in the way outlined in Sec. 3. The free energy and entropy per domain then follow from the partition function according to the standard formulas of statistical mechanics. If we want the entropy of a path of a fixed number of steps, or of a domain of a fixed perimeter, we must, in general, resort to the method of steepest descents in order to select the appropriate term from the generating function. However, this procedure is not necessary for the simple *path*-generating functions enumerated in Table III because they can be split into partial fractions and the coefficient of x^n written down in closed form, k times the logarithm of this coefficient being the entropy. For a long chain, only the factor corresponding to the *smallest* zero of the denominator

[17] C. M. Tchen, J. Chem. Phys. **20**, 214 (1952).

[18] H. C. Kramers and G. H. Wannier, Phys. Rev. **60**, 252, 263 (1941).

[19] G. S. Rushbrooke and H. I. Scoins, Phil. Mag. **42**, 582 (1951); **43**, 1276 (1952); Proc. Roy. Soc. (London) **A216**, 203 (1953).

[20] B. R. A. Nijboer and R. Fieschi, Physica **19**, 545 (1953).

$1-2\phi$ of the generating function contributes significantly to the entropy.

If we are discussing the entropy of a *domain* of known perimeter, we have either to use the method of steepest descents or to use appropriate expansions of the double integrals listed in Table I. Asymptotically, the entropy per leg of a domain is the same as that per leg of the corresponding path. This is a consequence of the fact that the domain generating function and the path generating function both become singular at the same value, x_c, of the selector variable. The path-generating function behaves in all cases like $(1-x/x_c)^{-1}$, which implies an entropy of $-nk \log x_c$ for a path of n steps. The domain generating function can also be expanded near $x=x_c$ in powers of x/x_c, but the coefficient of $(x/x_c)^n$ is now itself a function of n. The contribution to the entropy due to this coefficient is, in general, of the order of $k \log n$ and this can be neglected compared with nk for large n.

Wall *et al.*[10] have collected numerical evidence on the "probability of ring closure" for walks on certain types of lattice. Their "initial probability of ring closure" is simply the ratio of the number of domains of perimeter n to the total number of permitted paths of length n and is directly obtainable from our data. (The form of the connection of their "limiting probability of ring closure" with the present work has not yet been found.) Wall *et al.*[10] find empirically that, in two or three dimensions, the probability of closure on to the origin after n steps probably varies as n^{-2} for a restricted walk, in comparison with the known laws $n^{-\frac{1}{2}}$, n^{-1}, $n^{-\frac{3}{2}}$ for unrestricted walks on one-, two-, and three-dimensional lattices respectively. The corresponding information for our model can be obtained from the analytic behavior of the domain-generating function near its singularity already listed in Table I. Consider, for example, Model P. The number of paths is given by (constant independent of n) $\times x_c^{-n}$ (n large). The form $\log\delta$ for the domain-generating function implies that its principal term near the singularity behaves like $\log(x_c-x)$ so that the number of domains is given, by expanding the logarithm, as constant$\times x_c^{-n}/n$; so that, for model P, the probability of ring closure is proportional to $1/n$, as it is for the unrestricted model in two dimensions, a result that we might have anticipated.

For model P' the corresponding result is that the limiting probability of closure is $1/n^2$. The difference between P and P' arises from the different behavior of the domain-generating function, which in turn, arises from the form of the function $\phi(x)$, which now goes through a maximum exactly at $x=x_c$. The similar law found empirically for the two- and three-dimensional lattices by Wall *et al.*[10] seems to foreshadow a similar property of the domain-generating functions. Such behavior would be entirely consistent with what is already known about the Ising problem in two and three dimensions.

9. CONCLUSIONS

(a) Our results are in complete agreement with the conclusions from other work (see, for example, Yang and Lee)[8] that for the Ising plane square lattice, the boundary energy for a single domain persists right up to the "bulk" transition temperature, and that no lower transition temperature exists. This is also in accord with the known fact[21] that the spontaneous magnetization in the Ising model persists right up to the bulk transition temperature.

(b) Our work indicates that it may often be possible to take account of constraints (limiting possible types of path or domain) by leaving the *form* of the generating function unaffected, but transforming the *selector* variables, and possibly introducing into the generating function a multiplying factor. As an example of such a process, we may mention the Ising model, the removal of unwanted graphs being accomplished[4] by transforming the enumerating variable from x to $x(1-x^2)/(1+x^2)^2$. The mathematical significance of such transformations is far from clear, but there seems to be some deep reason for their existence.

(c) The comparison of the exact model R with models P' and Q, the two entirely different approximations to it, illustrates a state of affairs that is not particularly surprising from a mathematical point of view, but which does not yet seem to have occurred in any physical theory of a phase-transition. We have already noted that P' agrees with R up to and including six-sided domains, while Q agrees with R up to and including ten-sided ones. We have also noted that both models give a singularity of the correct *kind*, the generating function itself continuous, the derivative with respect to x being infinite. The precise analytic behavior of the two models differs slightly, Q not being quite right. We also notice that both Q and P' locate the correct transition temperature. Experience with the Ising model (for which some exact solutions are known) indicates that plausible approximations nearly always fail badly near the transition temperature, and that they sometimes also fail to give more than a few terms of the appropriate "high" and "low" temperature expressions.

Evidently, the mathematical reason why model P' is so good an approximation to R is the fortunate fact that "multiloop" paths tend to cancel out, while the success of model Q may be due to the fact that it considers explicitly a satisfactorily representative selection from the total number of possible domains that satisfy the constraints.

(d) These results may have a bearing on some recent investigations on approximations to the liquid distribution function,[19,20] since we have shown that an approximation may reproduce the "transition" behavior quite properly, and yet may fail to reproduce more than a few of the early terms of a series expansion. This differs from the behavior of many of the approximations in the

[21] C. N. Yang, Phys. Rev. 85, 808 (1952).

Ising model, which often reproduce the early terms well, but fail to predict the transition behavior properly.

(e) The elimination of paths which cross one another may lead to the breakdown of law (17), and a study of various generating functions has given us a clue to the possible mathematical reasons for such an occurrence, and for the closely related empirical fact that the "probability of path closure" becomes quite different, in the restricted case, even in three dimensions, from that for the unrestricted walk.

APPENDIX

We put on record the generating function that follows from the matrix P', a distinction now being made between diagonal and nondiagonal elements. Besides the application we have just made, this matrix may have others, more directly physical in type. In discussing problems connected with polymers, for example, it may be necessary not only to prevent paths from crossing one another, but to consider short-range correlation effects in addition. Thus, there is evidence that long-chain hydrocarbons up to about 30 atoms prefer to exist mainly in the "straightest" possible configuration, though a large number of "twisted" configurations are, in principle, available. This suggests some correlation between the directions of C—C valencies between neighboring "links," some configurations being energetically more probable than others, a situation that is theoretically quite likely.

The analytic effect of such a situation can be investigated in our model if we suppose that two successive links are more likely to form a "straight" than a "right-angle." We could take care of such a situation by supposing that in matrix P', the diagonal and nondiagonal elements contain respectively u and v instead of the single selector variable x, but that u and v are in a fixed ratio (associated with the postulated energy difference between the two types of configuration). The denominator of the generating function is then found to be

$$1+2u^2+u^4-4u^2v^2+4v^4 - (2u+2u^3-4v^2u)(\cos\omega_1+\cos\omega_2) +4(u^2-v^2)\cos\omega_1\cos\omega_2. \quad \text{(A1)}$$

The analytic behavior of this rather formidable expression does not appear to differ very much from that of the simpler one in (6), to which it reduces when $u=v$. It is readily found that this quantity will vanish only if $v=(1\pm u)/\sqrt{2}$ or $u=(1+2v^2)^{\frac{1}{2}}$.

These three cases arise from the possibilities

$$\cos\omega_1=\cos\omega_2=1,\ \cos\omega_1=\cos\omega_2=-1,\ \cos\omega_1\cdot\cos\omega_2=-1,$$

but only the first of these seems to be physically relevant. In all three of these cases we have the same feature (the function becoming a perfect square for $|\cos\omega_1|=|\cos\omega_2|=1$), which we have already noticed as the feature responsible for the differences between the "restricted" and "unrestricted" walk problem.