The algebraic construction of partition function zeros: universality and algebraic cycles

1987 J. Phys. A: Math. Gen. 20 3471

(http://iopscience.iop.org/0305-4470/20/11/048)

J. Phys. A: Math. Gen. 20 (1987) 3471-3493. Printed in the UK

The algebraic construction of partition function zeros: universality and algebraic cycles

D W Wood

Mathematics Department, University of Nottingham, Nottingham NG7 2RD, UK

Received 30 October 1986, in final form 26 January 1987

Abstract. This paper considers the one-parameter lattice models of classical statistical mechanics from a simple algebraic viewpoint. The way in which the limiting locus of partition function zeros emerges through a sequence of semi-infinite ($m \times \infty$) lattice sections is considered. Convergence to the limiting locus C_∞ is obtained through two sequences of algebraic curves. For all arbitrarily large but finite m the partition function per site is a branch of an algebraic function Λ_1 defined by an irreducible polynomial and possesses only a finite number of algebraic singular points. In the limit of $m \to \infty$ infinitely many branch points accumulate on C_∞; an algebraic basis to the universality hypothesis is presented in terms of the accumulation of branch points at the critical point. It is argued that for many two-dimensional lattice models the critical exponents α and ν will be of the form $\alpha = 2/s$ and $\nu = (s-1)/s$, where s is an integer $\geqslant 3$.

1. Introduction

A strong connection between the theory of algebraic functions and the mathematical construction of the critical behaviour which is exhibited by the lattice models of classical statistical mechanics in the thermodynamic limit has recently been suggested by the present author (Wood 1985). The purpose of the present paper and its companion paper (Wood *et al* 1987, hereafter referred to as II) is to explore this connection in detail. The present paper sets out the mathematical basis in the most general form which appears to be possible and as such addresses the set of all one-parameter lattice models. The general theory which is outlined here is illustrated in detail in II for a large number of the well known models of statistical mechanics.

2. The transfer matrix and irreducible polynomials

With respect to chosen boundary conditions, any classical lattice model with finite range interactions and discrete site variables can be defined by its transfer matrix. The boundary conditions are important in the present development to the extent of exploiting symmetry, and for this cyclic boundary conditions which permit maximal exploitation are assumed. By T_Ω we denote a transfer matrix of finite dimension Ω. T_Ω is established on a subgraph G of the lattice and with cyclic boundary conditions G is itself cyclically connected in $d-1$ dimensions. The lattice is generated by a one-dimensional translation of G in discrete steps. If G_1 and G_2 represent G translated through one such step the transfer matrix elements are defined by

$$T_\Omega(\alpha : \beta) = \exp -[\tfrac{1}{2}u(\alpha) + \tfrac{1}{2}u(\beta) + w(\alpha, \beta)] \tag{1}$$

and are indexed by α and β which are the discrete configurations of the classical site variables in G_1 and G_2, respectively. In (1), $u(\alpha)$ is that part of the reduced lattice Hamiltonian which is dependent solely upon α, and $w(\alpha, \beta)$ is that part of the reduced Hamiltonian dependent solely and simultaneously upon α and β.

The discrete configurations in G define a basis $\phi_1, \phi_2, \ldots, \phi_\Omega$ with respect to which T_Ω is defined. Thus for the scalar q-state Potts model $\Omega = q^m$, where the number of sites in G is m and $\phi_1, \phi_2, \ldots, \phi_\Omega$ are simply the q colourings of the subgraph G. Let m continue to denote the number of sites in G; the partition function of the $m \times n$ finite lattice is given by

$$Z_{mn}(z) = \sum_{i=1}^{\Omega} \lambda_i^n(z) \tag{2}$$

where the variable z is any variable in terms of which the transfer matrix can be expressed and $\lambda_i(z)$, $i = 1, \ldots, \Omega$, are the eigenvalues of T_Ω and hence roots of the characteristic equation

$$\lambda^\Omega + \psi_1(z)\lambda^{\Omega-1} + \psi_2(z)\lambda^{\Omega-2} + \ldots + \psi_\Omega(z) = 0 \tag{3}$$

where $\psi_k(z)$ are all rational functions of z. Equation (3) identifies all of the eigenvalues $\lambda_i(z)$ as algebraic functions. An algebraic function $y(z)$ is a function of z which satisfies a polynomial equation

$$P(y, z) = \phi_0(z)y^n + \phi_1(z)y^{n-1} + \ldots + \phi_n(z) = 0 \tag{4}$$

where the functions ϕ_k are analytic functions of z. Everywhere in the complex plane of z, y takes n determinations $y_1, y_2, \ldots, y_n$ which are the function elements or branches of the general analytic and multi-valued function $y(z)$. For $y(z)$ in (4) to be one such algebraic function it is necessary that the polynomial equation (4) be irreducible, where $P(y, z)$ does not factorise into two or more factors $P_1(y, z)P_2(y, z)$, where P_1 and P_2 are polynomials in the same class as $P(y, z)$. The eigenvalues of (3) belong to a specific class of algebraic functions, namely the class of functions for which ϕ_k in (4) are themselves all polynomials; this form can be developed for all transfer matrices by factorising out the largest inverse power of z from the matrix.

The eigenvalues of $T_\Omega(z)$ are never the branches of a single algebraic function since (3) is always reducible. This reducibility is a feature of all lattice models, the group of lattice rotations which map G into itself ensures that (3) is never irreducible, and it is always possible to block diagonalise the transfer matrix under a unitary transformation U which is itself independent of z, thus

$$U^+ T_\Omega(z) U = \bigoplus_k \tau_k(z) \tag{5}$$

where τ_k are matrices whose elements are polynomials in z. If a unitary transformation U can be found such that (5) constitutes a full factorisation of (3) into irreducible factors then a single algebraic function $\Lambda_k(z)$ can be assigned to each block, the branches of which are the eigenvalues of τ_k. A thorough description of this construction has been given by Ree and Chestnut (1966) and Runnels and Combs (1966) for two-dimensional hard core lattice-gas models; this construction generalises to any model in any dimension. In general this particular unitary transformation alone is unlikely to factorise (3) irreducibly for finite m because for most models the group of lattice rotations is simply a subgroup of a larger group under which the Hamiltonian is invariant. Consider, for example, any two-dimensional q-state Potts model; the

Hamiltonian is invariant to the group of m lattice rotations but also to the group of q colour rotations represented by the cyclic groups C_m and C_q, respectively. An ordering of the basis set $\phi_1, \phi_2, \ldots, \phi_\Omega$ is achieved by grouping subsets into equivalence classes which map into themselves under the chosen group. The group which minimises the number of equivalence classes will determine a transformation U which achieves a maximal factorisation of (3).

With $m=2$ in the q-state Potts model and the chosen group being C_2 the number of equivalence classes is $\frac{1}{2}q(q+1)$, whereas under the group C_q the number of equivalence classes is simply 2 and the transfer matrix takes the form

$$T=\begin{vmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{vmatrix} \tag{6}$$

where A_{11} and A_{22} are both cyclic matrices of dimension q and $q(q-1)$, respectively. The dimension of the largest block (τ_1 say) in (5) is 2; for the triangular lattice this matrix is

$$\tau_1(z)=\begin{bmatrix} z^2(z^4+q-1) & z(q-1)(z^3+z+q-2) \\ z(z^3+z+q-2) & qz^2+2(q-2)z+(q-2)^2 \end{bmatrix} \tag{7}$$

yielding a quadratic factor of the original characteristic equation (3) which was originally of degree q^2 in λ. Similarly for the simple quadratic lattice $\Lambda_1(z)$ is the two-valued algebraic function defined by the polynomial equation

$$\Lambda_1^2-[z^4+qz^2+2(q-2)z+q^2-3q+3]\Lambda_1+z^2(z-1)^2(z+q-1)^2=0 \tag{8}$$

which is irreducible. The remaining factors of the characteristic equation of (6) are simply linear factors and hence the remaining Λ_k are polynomials.

Finally we give an example in terms of the triplet Ising model on the triangular lattice (Baxter 1982). This model has the lattice Hamiltonian

$$H=-K\sum_{\Delta}\sigma_i\sigma_j\sigma_k \qquad (\sigma_i=\pm 1,\ i=1,2,\ldots,N) \tag{9}$$

where the three spin interactions are over all the elementary triangles of the lattice. Consider the case of a $3\times n$ lattice ($m=3$) under the group C_3 of lattice rotations. The four equivalence classes are

$$\begin{array}{ll} \{(+++)\} & \{(-++),(+-+),(++-)\} \\ \{(--+),(-+-),(+--)\} & \{(---)\} \end{array} \tag{10}$$

and $U=1\oplus C_3\oplus C_3\oplus 1$, where the largest matrix τ_1 is four dimensional. The group of lattice rotations is however not the only operation under which the transfer matrix elements are invariant. If we label the three site variables in G_1 and G_2 by t_1, t_2, t_3 and $\sigma_1, \sigma_2, \sigma_3$ we observe that $T(\alpha,\beta)$ is invariant to the operation $t_1, t_2, t_3 \to -t_1, -t_2, t_3$ and simultaneously $\sigma_1, \sigma_2, \sigma_3 \to \sigma_1, -\sigma_2, -\sigma_3$; the basis can be regrouped into two sets

$$\begin{array}{l} \{(---),(-++),(+-+),(++-)\} \\ \{(+++),(-+-),(+--),(--+)\} \end{array} \tag{11}$$

and T is factorised revealing that the previous quartic obtained from the largest block using (10) is reducible and (3) factorises to the form

$$\begin{aligned}&[\lambda^3+(2z^{-2}\sinh 4K)^3][\lambda^3-(2z^2\sinh 4K)^3][\lambda^2-2\lambda(3\cosh 2K+\cosh 6K)\\ &\quad +6(\cosh 8K-1)] \qquad (z=e^K).\end{aligned} \tag{12}$$

Some interesting examples of the block diagonalisation (5) for the vertex model form of the q-state Potts model have recently been given by Martin (1986) (see II, § 3).

For most lattice model problems in statistical mechanics the transfer matrix is a strictly positive matrix for $z>0$, where $T(\alpha,\beta)>0$, $z>0$ with the exception of hard core type models but in such cases there exists an integer k for which $T_\Omega^k(z)$ is a positive matrix. Hence for such lattice models the theorem of Perron and Frobenius applies to $T_\Omega(z)$ on the positive real axis, namely that the eigenvalue of maximum modulus is positive and simple for all $z>0$. This eigenvalue belongs to one of the irreducible factors of (3); we will denote the algebraic function to which this eigenvalue belongs as $\Lambda_1(z)$ and the block which defines Λ_1 as $\tau_1(z)$. The elements of $\tau_1(z)$ can always be expressed as polynomials in z (if the factorisation of T_Ω has been performed) with positive coefficients, and so clearly the Perron and Frobenius theorem applies to $\tau_1(z)$ alone. In fact $\tau_1(z)$ is the only matrix in the set τ_k which is positive for $z>0$. This alone suffices to identify $\tau_1(z)$ as the matrix which generates the algebraic function $\Lambda_1(z)$ which determines the maximum eigenvalue $\Lambda_1^+(z)$, $z>0$, where

$$\lim_{n\to\infty}(1/n)\ln Z_{mn}=\ln\Lambda_1^+(z). \tag{13}$$

We can summarise our results as follows.

(i) The characteristic equation of $T_\Omega(z)$ for all model problems defined upon a lattice factorises into irreducible polynomials.

(ii) The eigenvalue Λ_1^+ for $z>0$ is at all times a branch of one such polynomial factor, $\Lambda_1(z)$, and remains distinct from all the other branches of Λ_1 for $z>0$.

(iii) For the most part the factorisation of (3) into irreducible factors can be accomplished by a single unitary transformation (5), defined by the largest combined group of operations in both lattice space and spin space under which the lattice Hamiltonian remains invariant.

(iv) The eigenvectors of $T_\Omega(z)$ which belong to the branches of $\Lambda_1(z)$ all transform identically into themselves under the group which defines U.

3. Branch points, connection curves and more branch points

In § 4 we discuss the manner in which the algebraic construction of the limiting locus of partition function zeros proceeds in the limit of $m\to\infty$. In this a spectacular geometry of algebraic curves emerges. The stochastic build-up of a Julia set for hierarchical models (Derrida *et al* 1983) is replaced by an intricate system of smooth algebraic curves which converge onto the limiting locus. The construction of this locus involves an interplay between the limiting locus of the $m\times\infty$ system itself and the locus of zeros generated by the block $\tau_1(z)$ in (5) considered as though it were a partition function in its own right. The former is pathological in the limit of $m\to\infty$ and the latter is smooth and well defined; both converge to the locus of partition function zeros in the thermodynamic limit of m, $n\to\infty$. All of these algebraic curves are generated by the polynomial resolvent introduced by Wood (1985) which defines a cut in the z plane where the eigenvalues of T_Ω are everywhere distinct in modulus in the cut plane. Of particular importance are the cuts generated by the single block $\tau_1(z)$ in which the algebraic singular points (branch points) of Λ_1 are linked by a network of algebraic curves exhibiting a rich geometrical structure. Although the theory of algebraic functions is a highly developed area (Fuchs and Shabat 1961, Goursat 1916, Appell and

Goursat 1929, Ahlfors 1953) the author has found little in the mathematical literature relating to these geometrical features. In this section we discuss these novel aspects with reference to algebraic functions in general before considering the algebraic functions generated by the transfer matrices of statistical mechanics.

We consider an algebraic function Λ with n branches $\lambda_1, \lambda_2, \ldots, \lambda_n$ defined by the irreducible polynomial equation

$$F(\Lambda, z) = \Lambda^n + \phi_1(z)\Lambda^{n-1} + \phi_2(z)\Lambda^{n-2} + \ldots + \phi_n(z) = 0 \tag{14}$$

where $\phi_k(z)$ are polynomials in z. Since $\phi_0 = 1$, Λ has a finite number of algebraic singular points (branch points) including the circle at infinity, and no poles at finite values of $|z|$; it is regular everywhere in the z plane apart from these points. The branch points are arbitrary in number. Since $\phi_k(z)$ are polynomials the branch points are easy to locate formally; they are the roots of a polynomial (the discriminant)

$$\Delta(z) = 0$$

where $\Delta(z)$ is the resolvent between the two polynomials

$$F(\Lambda) = 0 \tag{15}$$

and

$$\partial F(\Lambda)/\partial\Lambda = 0. \tag{16}$$

The degree of $\Delta(z)$ depends upon the polynomials $\phi_k(z)$; when the coefficients of $\phi_k(z)$ are all real then the branch points of Λ occur in complex conjugate pairs. Well known forms for $\Delta(z)$ are the determinantal forms of Sylvester and Bezout (Turnbull 1952, Littlewood 1958).

Let z_i, $i = 1, 2, \ldots$, be the branch points of Λ, and consider the set L of points in the z plane where any pair λ_i and λ_j of the n branches are equal in modulus. Thus on L

$$\lambda_i = \lambda_j \, e^{i\phi} \tag{17}$$

for some i and j; clearly z_i belong to L. The set L is in fact traced out by smooth algebraic curves which are themselves branches of an algebraic function Z. We shall refer to these curves as connection curves, since they define a set of connections between the branch points z_i. The connection curves are defined formally in terms of the resolvent between the two polynomials

$$F(\Lambda) = 0 \qquad F(h\Lambda) = 0 \qquad (|h| = 1). \tag{18}$$

In determinantal form this is specifically given by the polynomial equation (Littlewood 1958)

$$\begin{vmatrix}
1 & \phi_1 & \phi_2 & \ldots & \phi_n & & & & & & \\
 & 1 & \phi_1 & \phi_2 & \ldots & \phi_n & & & & & \\
 & & & & & & & & & & \\
 & & & & & & 1 & \phi_1 & \phi_2 & \ldots & \phi_n \\
h^n & h^{n-1}\phi_1 & h^{n-2}\phi_2 & \ldots & \phi_n & & & & & & \\
 & h^n & h^{n-1}\phi_1 & h^{n-2}\phi_2 & \ldots & \phi_n & & & & & \\
 & & & & & & & & & & \\
 & & & & & & h^n & h^{n-1}\phi_1 & h^{n-2}\phi_2 & \ldots & \phi_n
\end{vmatrix} = 0 \tag{19}$$

and defines the branches $z_1(h), z_2(h), \ldots, z_s(h)$ of an algebraic function Z defined by the expansion of the determinant yielding

$$(1-h)^n[\phi_0(h)Z^s + \phi_1(h)Z^{s-1} + \ldots + \phi_s(h)] \equiv (1-h)^n R(z, h) = 0 \tag{20}$$

where $\phi_i(h)$ are polynomials in h. Strictly we should first assure ourselves that (20) is itself irreducible before we can define a single algebraic function Z. At $h=1$ (20) is an identity since the two polynomials in (18) are identical and the roots of $R(z,1)$ determine the branch points of Λ.

Consider the polynomial equation

$$R(h,z)R(h^*z)=0 \tag{21}$$

invariant under the transformation $h\to h^{-1}$. The real coefficients in (21) are mapped out in terms of the variable

$$w=\cos\phi=\tfrac{1}{2}(h+h^{-1}) \tag{22}$$

and we need consider only the domain $0\leqslant\phi\leqslant\pi$. As ϕ increases from 0 to π the branches $z_k(\phi)$ trace out complex conjugate curves which emerge from the branch points $z_k(1)$ and are completed at $\phi=\pi$. With the circle at infinity included as a possible branch point the curves traced out by $z_k(\phi)$ define a set of connections on the branch points $z_k(1)$, which is the set L and which we now refer to as connection curves.

The branches λ_i are the function elements of Λ, which has a finite number of algebraic singular points. Without any restrictions imposed on a path from A to B we can always find one path where the branch $\lambda_1(z)$, say, at A arrives at B with any one of the values of the $n-1$ other branches. Single valuedness in the branches $\lambda_i(z)$ themselves can always be achieved in the cut plane. The set of connection curves L above is such a construction where at all points in the z plane cut by the connection curves the roots of (14) are distinct in modulus. Not surprisingly the geometry of the connection curves L is extremely intricate and highly structured. In relation to the algebraic functions generated by the transfer matrices of statistical mechanics the z plane cut by the connection curves is important to us since it reveals paths in the plane where the branch of $\tau_1(z)$, which is maximum in modulus, remains maximum in modulus along all points in the path.

We now recall the analytic development of the branches $\lambda_i(z)$ about an algebraic singular point $z_k(1)$. At a branch point $z_k(1)$ a number of the roots λ_i are equal. Consider the multiplicity of these roots to be m_k at $z_k(1)$. As is well known (Fuchs and Shabat 1961) the m_k roots group themselves into a cycle structure $p_1,p_2,\ldots$, where

$$m_k=p_1+p_2+\ldots \tag{23}$$

and p_1 of the m_k branches which become equal at $z_k(1)$ permute into each other on a path in the z plane which encircles $z_k(1)$ and does not enclose another singular point. Similarly, p_2 other function elements from amongst the same m_k branches permute into each other, and so on. In such a grouping of branches the roots which permute into each other are said to belong to a cycle and p_i is the cycle number of this group. The branches of the same cyclic system with cycle number p, say, at a branch point $z_k(1)$ can be represented by a unique analytic development about the branch point (here assumed to be an algebraic zero), in the form

$$\lambda_i=\lambda_i(z_k(1))+\sum_r a_r\varepsilon^{r/p} \qquad i=1,2,\ldots,p \tag{24}$$

where $\varepsilon^{1/p}$ takes all of its p determinations, a_r are constants and $\varepsilon=(z-z_k(1))$.

We now consider what can be said about the geometry of the connection curves L. Firstly, there are two cases to be identified in relation to the nature of connection curves in the neighbourhood of a branch point $w\sim 1$.

(a) The number of branch points $s =$ the degree of $R(z, w)$ as a polynomial in z.
(b) The number of branch points $s <$ the degree of $R(z, w)$.

In case (a) one connection curve emerges from each branch point. In case (b) the discriminant $R(z, 1)$ has itself multiple roots and some of the branch points of Λ correspond to non-analytic points of the branches $z_k(w)$, and so $w = 1$ is also a branch point of Z in (20). Clearly at those branch points more than a single curve emerges. Suppose p is a cycle number of this branch point. Applying (24) to the branches of (20) we have

$$z_j(w) = z_j(1) + a_1\varepsilon^{1/p}\omega_j + \dots \qquad j = 1, 2, \dots, p \tag{25}$$

where $\varepsilon = 1 - w \geq 0$ and ω_j are the p roots of unity. Equation (25) describes the nature of the connection curves in the neighbourhood of this type of branch point $z_k(1)$, since $\varepsilon \geq 0, p$ curves emerge from $z_k(1)$ asymptotically along curves which subtend angles of $2\pi/p$ with each other.

The other branch points of Z in the domain of $|h| = 1$ also play an important role in the way in which the connection curves connect the branch points z_k of Λ. These points are intersection points of the connection curves as they are traced out under increasing ϕ, corresponding to values of ϕ or w where the resolvent $R(z, w)$ has itself multiple roots. The multiple roots of $R(z, w)$ in the domain $|h| = 1$ are the intersection points of L. If w_c ($\neq 1$) is a branch point of Z and p is a cycle number associated with this branch point, then in the neighbourhood of $w_c p$ branches of Z have the form

$$z_j(w) = z_j(w_c) + a_1\varepsilon^{1/p}\omega_j + \dots \qquad j = 1, 2, \dots, p \tag{26}$$

where $\varepsilon = w_c - w$ and is real. For $p = 2$ a pair of connection curves cross orthogonally and $z_j(w_c)$ is in general the meeting point of $2p$ curves asymptotically subtending angles of π/p with each other. All aspects of the above geometry appear in the applications to statistical mechanics included in II, but it will be helpful here to include a few illustrations based upon simple polynomials of the type (14), and indeed the simplest of examples unfolds a striking geometry.

Consider a two-valued function Λ_n defined by

$$\Lambda_n^2 - 2\Lambda_n + z^n = 0. \tag{27}$$

This has two branches

$$\lambda_{1,n} = 1 + (1 - z^n)^{1/2} \qquad \lambda_{2,n} = 1 - (1 - z^n)^{1/2} \tag{28}$$

where the positive and negative signs originate in the two determinations of $(1 - z^n)^{1/2}$. For any quadratic equation

$$\Lambda^2 + b\Lambda + c = 0 \tag{29}$$

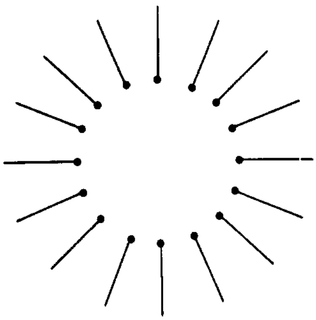
the connection curves are traced out by the solutions to the equation

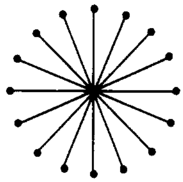
$$b^2(1 + t^2) = 4c \qquad (t = \tan\phi/2) \tag{30}$$

which is the form taken by the resolvent $R(z, w)$ using (21). The branch points of Λ_n in (27) are the n roots of unity ω_j and hence the connection curves are

$$z_i(\phi) = (1 + t^2)^{1/n}\omega_j \qquad j = 1, 2, \dots, n \tag{31}$$

namely the spokes shown in figure 1(a) which connect each branch point to the circle at infinity. Here we have a case where in the limit of $n \to \infty$ the branch points become





(a) (b)

Figure 1. (a) The connection curves of (27), (b) the connection curves of (32).

dense on the unit circle with a uniform distribution, and also in this limit the connection curves become dense in $|z|>1$. Redefining Λ_n to be the roots of

$$\Lambda_n^2-2z^n\Lambda_n+1=0 \tag{32}$$

produces branch points at the $2n$ roots of unity and the resolvent equation now becomes

$$z^{2n}(1+t^2)=1 \tag{33}$$

which itself has a branch point in the limit of $t\to\infty$ where $z=0$ is a root of multiplicity $2n$. The connection curves are shown in figure 1(b). They are the spokes now interior to the unit circle and $z=0$ is a meeting point of these curves which constitute n straight-line segments connecting the points $\pm\omega_j$, $j=1,2,\ldots,2n$.

Examples of the above type can be designed without recourse to a polynomial equation. If an analytic function $\Lambda(z)$ has only n distinct values for each z, and if it has in the whole plane (including the point at ∞) only algebraic singular points, then the n determinations of Λ are the roots of a polynomial of degree n in Λ whose coefficients are rational functions of z. Consider the algebraic function $\Lambda_{m,n}$ at particular values of m and n given by

$$\Lambda_{m,n}=(1+z)^m\pm(1-z^n)^{1/2} \qquad n \text{ odd} \tag{34}$$

where we have maintained the branch points of $\Lambda_{m,n}$ as the n roots of unity but have simultaneously created algebraic singularities in the branches $z_i(t)$ of the resolvent equation (of degree $2m$ $(2m>n)$ or n $(2m<n)$)

$$(z^n-1)=(1+z)^{2m}t^2 \qquad (t=\tan\phi/2) \tag{35}$$

at real values of t. Transforming (35) under $z=1/v$ yields

$$v^{2m-n}(1-v^n)=(1+v)^{2m}t^2 \qquad 2m>n \tag{36}$$

showing that at $\phi=0$ (the origin of the connection curves, and a branch point of $\Lambda_{m,n}$) the point at infinity is a root of multiplicity $2m-n$. Thus $2m-n$ connection curves emerge from the circle at ∞, and do so in the asymptotic directions equal to those defined by the $2m-n$ roots of unity. Similarly, transforming (35) under $t\to 1/t$ yields the fact that $z=-1$ is a root of the resolvent of multiplicity $2m$, and hence $z=-1$ is a meeting point of all the connection curves which leave a branch point. When $2m>n$ there is also a root of the resolvent of multiplicity 2 on the real axis somewhere between 1 and 4. It remains to discover the geometry of these connections. Consider the cases

$n=3$, $m=1,2,4\ldots$. With $m=1$ there is no branch point at infinity and $z=-1$ is the meeting of connection curves emanating from the two complex roots of unity—the situation is shown schematically in figure 2(a). For $m>1$ the cycle number of the branches $z_i(t)$ at $z=-1$ is $2m$. Here $a_1=0$ and $a_2\neq 0$ whence the branches z_j in the neighbourhood of $z=-1$ are of the form

$$z_j=-1+a_2\varepsilon^{1/m}w_j^2+\ldots \qquad j=1,2,\ldots,2m,\ \varepsilon>0. \tag{37}$$

Thus, with $m=2$, the branches follow the curves shown in figure 2(b). For $m=4$, the curves intersect orthogonally as shown in figure 2(c).

The above analysis of the resolvent (20) gives some classification of the geometry of the connection curves but it does not indicate which of the branch points of the function $\Lambda(z)$ in (14) become connected. A feature of the polynomial generated in the characteristic equation of the block $\tau_1(z)$ in (4) is that Λ_1^+ has no algebraic singular points at ∞ (see § 4). The most common feature of the connection curves in this case is that they appear to constitute a set of connections between pairs of branch points. The question arises as to what can be said about such pairs which are connected.

Consider a point on a connection curve traced out by a branch $z_i(w)$ of the resolvent. At such a point there exists at least two branches λ_i and λ_j which are equal in modulus. It follows that λ_i and λ_j can be written in the form

$$\lambda_i=uy \qquad \lambda_j=uy^* \tag{38}$$

where $|u|=1$; thus λ_i and λ_j are roots of a quadratic factor of the form

$$\Lambda^2+bu\Lambda+u^2c \tag{39}$$

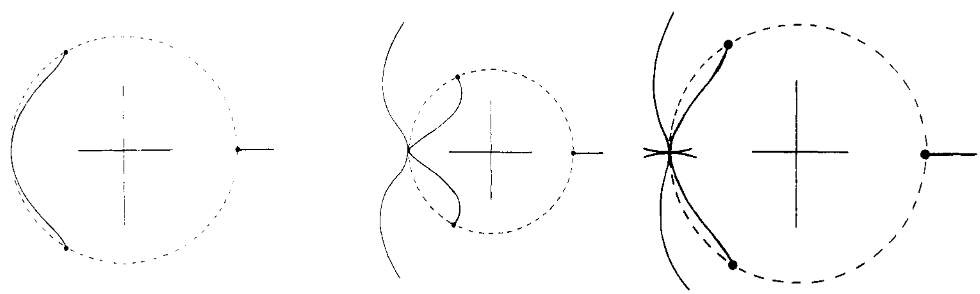
and there exists therefore a real quadratic factor

$$y^2+by+c=0 \tag{40}$$

with roots y_1 and y_2 which are complex conjugate pairs when $z_i(w)$ is a point on a connection curve. Now $w(=\cos\phi)$ is a parametrisation of this connection curve and everywhere on it away from a branch point (40) takes the form

$$y^2+b(w)y+c(w)=0 \tag{41}$$

where the real coefficients are smoothly varying functions of w. The roots of (41) yield the two branches in (38) and at the point $w=1$ where y_1 and y_2 are both equal and real we move into a branch point of λ_i and λ_j. We see that the branch points of (41) are linked by the smooth connection curve $z_i(w)$, and so if two branch points of $\Lambda(z)$



(a) (b) (c)

Figure 2. The connection curves of (34), (a) $m=1$, $n=3$, (b) $m=2$, $n=3$, (c) $m=4$, $n=3$.

are connected by a smooth curve they are both the branch points of a specific pair of branches of $\Lambda(z)$. Equation (21) provides a natural extension of a connection curve through its two branch points, assuming for the moment that $z_i(w)$ connects just two branch points at complex values of z. If we extend the domain of w from $-1 \leq w \leq 1$ to the whole real line of w, then for $|w| > 1$ (which is equivalent to the hyperbolic extension $\phi \rightarrow \pm i\phi$ in (17)) the connection curve continues smoothly through each of the branch points and may continue on a trajectory which intercepts the real axis corresponding to the algebraic function Z of the resolvent having a branch point at a pure imaginary value of ϕ. This hyperbolic extension is of great importance in statistical mechanics and is discussed in II (§ 2).

For cases where the polynomials $\phi_k(z)$ in (14) have some simple algebraic structure of their own the entire equation could be linked to a polynomial with real coefficients everywhere in the z plane on a smooth curve $z(w)$. In this case a large portion of the branch points of Λ are highly structured since large numbers of them probably lie on a single smooth curve. This phenomenon is of great importance in statistical mechanics where the symmetries of the transfer matrix can lead to precisely this global effect upon the characteristic polynomial of $\tau_1(z)$ or indeed of the whole transfer matrix T_Ω. A striking example of this is the self-duality symmetry which is a property of the partition function of numerous two-dimensional models. Details of this are given in II; here we note that models with self-dual symmetry are such that in the large system limit of arbitrarily large m the branch points of the eigenvalues of $\tau_1(z)$ are arbitrarily many in number and a subset of them lie on a single invariant connection curve. It is precisely this phenomenon which is the algebraic origin of the original Yang–Lee circle theorem for all Ising models (Yang and Lee 1952, see II, § 8).

In this section we have discussed the algebraic and geometric structure of a cut L in the plane of the variable z in terms of which a many-valued function $\Lambda(z)$ defined by (14) is mapped out. In this cut plane the branches of Λ are regular and everywhere distinct in modulus; the cut L is a set of algebraic curves connecting the branch points of Λ. These connection curves are themselves the branches of a single algebraic function defined by the resolvent polynomial equation (19) previously introduced by Wood (1985). The geometry of the connection curves is characterised by their form at the branch points of Λ and the nature of their own intersection points which are determined by the branch points of the algebraic function (20) defined by the resolvent and their associated algebraic cycle structure. Any arc of a connection curve connecting a pair of branch points defines a set of points in the z plane where the defining polynomial of Λ (14) contains a multiple of a real quadratic factor. In the following section we trace the relationship between the locus of partition function zeros and the connection curves generated by the block $\tau_1(z)$ of the transfer matrix through its characteristic equation, for a general one-parameter lattice model in statistical mechanics.

4. The limiting locus of partition function zeros

We now consider in general terms the manner in which the partition function (2) constructs an asymptotic distribution of zeros in the z plane in the limit of $m \rightarrow \infty$. In this there are two aspects which must be clearly distinguished at the outset. These are the asymptotic distribution of zeros of the semi-infinite lattice obtained in the limit of $n \rightarrow \infty$, m finite, and the asymptotic locus obtained in the full thermodynamic limit of $m \rightarrow \infty$ and $n \rightarrow \infty$. In the former case we denote the asymptotic locus of zeros by C_m

which becomes C_∞ in the thermodynamic limit of $m \to \infty$. The purpose of this section is to present in general terms the relationship between the loci of the sequence C_m ($m = 2, 3, \ldots$), the connection curves formed by the eigenvalues of $\tau_1(z)$ (denoted by the algebraic function Λ_1) and the asymptotic locus of zeros C_∞. In this it emerges that the connection curves of Λ_1 are of greater interest and significance than C_m in relating a study of partition function zeros to critical behaviour. The branch point structure of Λ_1 and the nature of the connections between them provide, respectively, a view of the way in which the critical singularities emerge as the limit of an infinite sequence of algebraic singular points, and a new practical way of obtaining accurate determinations of both critical points and the non-physical singularities of the limiting partition function per site.

The partition function $Z_{mn}(z)$ is the symmetric function of all of the eigenvalues of the transfer matrix $T_\Omega(z)$ given by (2). Consider the eigenvalues $\lambda_i(z)$ of $T_\Omega(z)$ ordered in modulus

$$|\lambda_1| \geqslant |\lambda_2| \geqslant \ldots \geqslant |\lambda_\Omega| \tag{42}$$

and let C_m be the subset of the z plane where $Z_{mn}(z)$ is asymptotically zero in the limit of $n \to \infty$, by which when $|\lambda_1| \geqslant 1$ we shall mean

$$\lim_{n\to\infty} \frac{Z_{mn}(z)}{|\lambda_1(z)|^n} = 0. \tag{43}$$

We arrive at the following mechanisms by which z is in the set C_m.

(i) If $|\lambda_1(z)| < 1$, then $z \in C_m$.

(ii) If λ_1 is distinct in modulus and $|\lambda_1(z)| \geqslant 1$ then $z \notin C_m$.

(iii) If $\lambda_1 \neq \lambda_2$ and $|\lambda_1(z)| = |\lambda_2(z)| \geqslant 1$ then $z \in C_m$.

Cases (i) and (ii) are obvious; in case (iii)

$$\frac{Z_{mn}(z)}{\lambda_1^n(z)} = 1 + \mathrm{e}^{in\phi} + \sum_{k=3} (\lambda_k/\lambda_1)^n \tag{44}$$

and

$$\lambda_1/\lambda_2 = \mathrm{e}^{i\phi}. \tag{45}$$

For all ϕ which are rational multiples $(2j+1)/n$ of π, $1+\mathrm{e}^{in\phi}$ is zero infinitely often in the limit of $n \to \infty$ and, since $|\lambda_k/\lambda_1| < 1$, $k \geqslant 3$, (43) is satisfied (see also Martin 1986). When ϕ is an irrational multiple of π it follows from Kronecker's theorem (Hardy and Wright 1960) that $|1+\mathrm{e}^{in\phi}| < \varepsilon$ for any positive ε at infinitely many values of n, and the partition function can be made as close to zero as we please.

It has frequently been observed in the literature that C_∞ often appears to be a smooth curve for examples of one-parameter models which have been investigated either analytically or numerically. We can see in (45) the mathematical reason why this is to be expected for C_m. Conditions (i) and (iii) reveal two possible geometries for C_m. Condition (i) allows for the possibility that C_m is dense inside a bounded region of the z plane even for one-parameter models, although this does not appear to be an option taken up by such models. Condition (iii) and equation (45), however, identify C_m to be made up of algebraic curves. These curves are a subset of the algebraic curves $z_i(w)$ generated by the resolvent equation (20) in the form of (21) formed on the characteristic polynomial of the whole of the transfer matrix $T_\Omega(z)$. We shall denote this resolvent by $R_\Omega(z, w)$.

In (iii) the branches λ_1 and λ_2 can be viewed as cooperating in determining part of C_m. However more than two branches can cooperate in this way and a plethora of extensions to (iii) unfolds. Thus we could extend (iii) by adding the following.

(iv) If $\lambda_1 \neq \lambda_2 \neq \lambda_3$ and $|\lambda_1| = |\lambda_2| = |\lambda_3|$ and

$$\lambda_1 : \lambda_2 : \lambda_3 = 1 : \mathrm{e}^{\mathrm{i}\phi} : \mathrm{e}^{-\mathrm{i}\phi} \tag{46}$$

then $z \in C_m$ for all ϕ.

This condition adds no new regions to the z plane than those already found from the resolvent $R_\Omega(z, w)$, but we see that the subset of these curves which form C_m are not exhaustively classified by (iii). It is to be expected however that condition (iii) would be the most common mechanism in generating the asymptotic zeros of $Z_{mn}(z)$ in the limit of $n \to \infty$. In a recent paper, Martin (1986) has identified condition (iii) in a study of the square lattice q-state Potts model using a novel factorisation (5). The claim is that finite $m \times n$ lattice partition functions can give a good image of the limiting distribution C_m, but no examples of C_m itself are presented (see II, § 3). It is important to understand that at this stage the algebraic curves which form C_m cannot yet be identified as connection curves of the type defined in § 3. The curves generated by $R_\Omega(z, w)$ will be of two types: (a) sets of connection curves C_m^k generated by the irreducible characteristic polynomials of each block $\tau_k(z)$ in (5) which form a linking of the branch points of Λ_k, and (b) the sets of curves C_m^{kj} tracing out paths in the z plane where branches of Λ_k and Λ_j ($k \neq j$) are equal in modulus. We will call such curves cross block curves since they are not related to the branch point structure of either Λ_k or Λ_j.

The labyrinthine nature of the geometry which unfolds can only really be understood at the outset by example. Here we take two simple examples; many others are given in II. As a very simple example we consider the nearest-neighbour Ising model on a $2 \times n$ section of the quadratic lattice, where the 4×4 transfer matrix has three blocks which define one two-valued function and two single-valued functions. These are

$$\Lambda_1^2 + 4(1+s^2)\Lambda_1 + 4s^2 = 0 \qquad (s = \sinh 2K) \tag{47}$$

$$\Lambda_2 = z^2 - 1 \tag{48}$$

and

$$\Lambda_3 = 1 - z^{-2} \qquad (z = \mathrm{e}^{2K}). \tag{49}$$

The curves traced out by the resolvent $R_\Omega(z, w)$ (here $\Omega = 4$) are shown in figure 3 in the z plane. The connection curve of Λ_1 lies on the unit circle $|s| = 1$; this can be seen immediately since on writing $\Lambda_1 = sy$,

$$y^2 + 4(s + s^{-1})y + 4 = 0. \tag{50}$$

This is an example of the quadratic factor (40), since on $|s| = 1$ (50) is a real quadratic factor and the connection curves are the arc of $s = \mathrm{e}^{\mathrm{i}\theta}$ connecting $\theta = \pi/3$ and $2\pi/3$ and its complex conjugate (Wood 1985). In the z plane these are the four arcs shown in figure 3 on the circles $z = \pm 1 + \sqrt{2}\,\mathrm{e}^{\mathrm{i}\theta}$. The remaining curves in figure 3 are all cross block curves, the unit circle is C_2^{23} and the curves C_2^{12} and C_2^{13} are reciprocal curves; in fact viewed as a whole the curve in figure 3 is self-reciprocal.

Having constructed the set of curves generated by the resolvent $R_\Omega(z, w)$ the question arises as to which of these is C_2. In this the nature of the Gibbs distribution allows for a powerful observation. In the limit of zero temperature which we take to be the

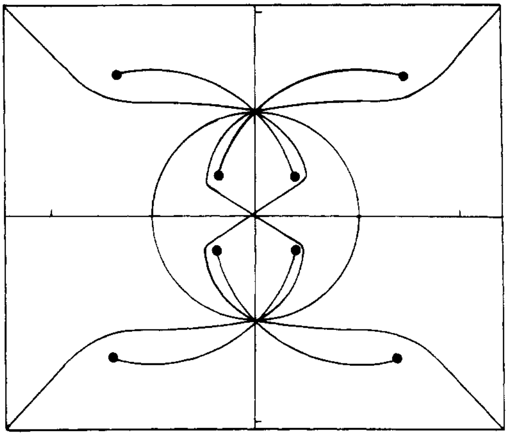


Figure 3. The connection curves and cross block curves traced out by $R_\Omega(z, w)$ $(z = e^{2K})$ obtained from the transfer matrix of the nearest-neighbour Ising model on a $2 \times n$ strip of the quadratic lattice; ● denotes a branch point.

limit of $|z| \to \infty$ any thermodynamic system is frozen into its ground state, namely the configurations in $\phi_1 \ldots \phi_\Omega$ which are of minimum energy. For models which exhibit any form of phase equilibrium the ground state is necessarily degenerate and the eigenvalues of $T_\Omega(z)$ are asymptotically degenerate in the limit $|z| \to \infty$, where λ_1 in (42) is l-fold degenerate for any model with l coexisting ground states. Condition (iii) above ensures that the circle at infinity is part of C_m. It is immediately obvious that the curves of $R_\Omega(z, w)$ belonging to C_m and meeting on the circle at infinity cannot be connection curves for if they were this circle would be a branch point at infinity and the thermodynamic functions would all have an algebraic singular point at zero temperature, which is clearly impossible. Therefore the curves belonging to C_m and meeting at infinity are all cross block curves and there always exists a circle $|z| = \rho$ where, for $|z| > \rho$, C_m is composed only of cross block curves. Here we have a hint that C_m is composed of such curves (unless and until an intersection with a connection curve occurs), namely the continuation of this set of smooth curves lying outside $|z| = \rho$. This assumption has an important geometrical consequence for the connection curves C^1_m, which follows from the theorem of Perron and Frobenius.

The z plane cut by the curves C^1_m and C^{1k}_m is a plane in which the branches of Λ_1 are all distinct in modulus from all other eigenvalues of the transfer matrix. We know that Λ_1^+ is distinct and maximum in modulus on the positive real axis of z. It follows from this that if the connection curves in C^1_m are not part of C_m they must be completely fenced off from the real axis by curves from the set C^{1k}_m. That is, starting at a point A on the real positive axis of z there is no path in the z plane connecting A to the connection curves of Λ_1 involving the branch Λ_1^+ which does not intersect a cross block curve. If there were such a path then those parts of C^1_m which could be reached in this way would satisfy (iii) and hence be part of C_m. Such an effect is strikingly illustrated in figure 3 where C^{12}_2 and C^{13}_3 have completely isolated the connection curve C^1_2 from the whole of the real axis. It follows from the above argument that C_2 in this case is made up of C^{12}_2 and C^{13}_2; C^{23}_2 is redundant.

As a second and slightly more complicated example we consider the same $2 \times \infty$ quadratic lattice for an Ising model with next-nearest-neighbour interactions of equal strength to the nearest-neighbour interactions and in terms of the variable $z = e^{4K}$.

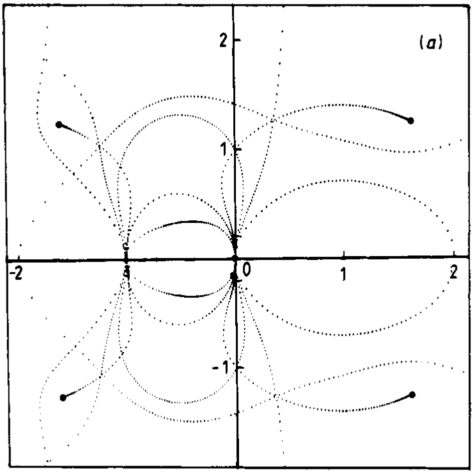
Again $T_4(z)$ has three irreducible blocks and the eigenvalues are given by

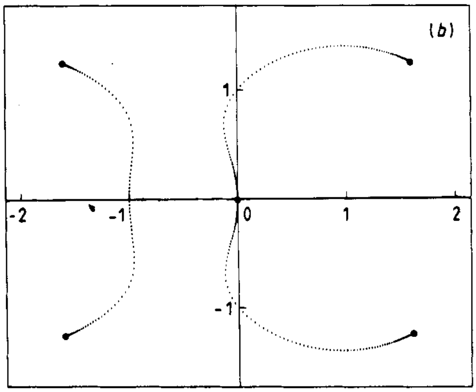
$$\Lambda_1^2 - \Lambda_1(z^3 + z + 2) + (z^4 + z^3 - 4z^2 + z + 1) = 0 \tag{51}$$

$$\Lambda_2 = z^3 - 1 \tag{52}$$

$$\Lambda_3 = 1 - z. \tag{53}$$

$R_4(z, w)$ traces out the intricate system of curves shown in figure 4(a) which are decomposed in figures 4(b) and (c) into the connection curves of Λ_1 and the cross block curves C_2^{ij}, respectively. We see again that the connection curve C_2^1 is isolated from the real positive axis by a fence formed by C_2^{12} but now involving the trisection point which lies on the C_2^1 curve. Curves such as C_2^{13} which intersect the real positive axis can of course be immediately eliminated from C_2. This curve is formed by the branch of Λ_1 which is not Λ_1^+. We can readily determine which eigenvalue of the transfer matrix is of maximum modulus in the various regions of the z plane. This is illustrated in figure 5(b), from which we confirm that the curves of figure 5(a) are those which form C_2. Here C_2 is the smooth extension of the curves C_2^{12} connected to the circle at infinity but now only up to an intersection point which is also a point on C_2^1, and C_2^1 is partially isolated from the positive axis by C_2.





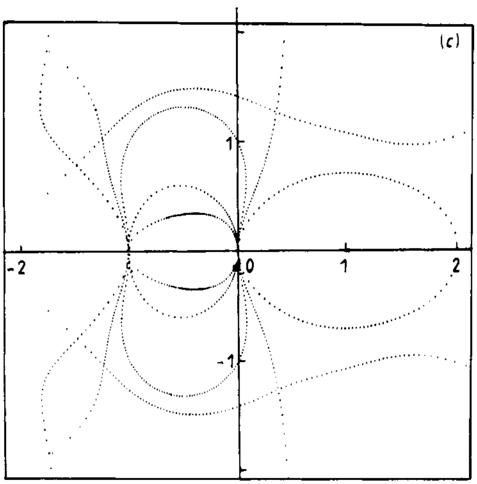


Figure 4. (a) The connection curves and cross block curves traced out by $R_\Omega(z, w)$ ($z = e^{4K}$) obtained from the transfer matrix of the Ising model with equal nearest-neighbour and next-nearest-neighbour interactions on a $2 \times n$ strip of the quadratic lattice; ● denotes a branch point. (b) The connection curves in (a); ● denotes a branch point. (c) The cross block curves in (a).

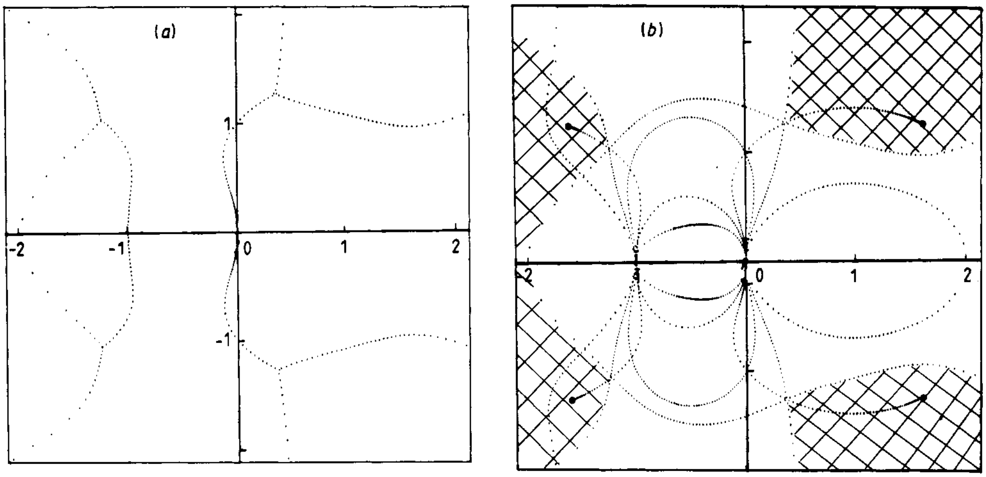


Figure 5. (a) The curves in figure 4(a) which form the limiting locus of zeros C_2. (b) The connection curves in figure 4(a) in regions where an eigenvalue of $T_\Omega(z)$ other than a branch of Λ_1 is maximum in modulus. In the shaded region Λ_2 is maximum in modulus.

Clearly, as m increases a geometry of almost unimaginable complexity unfolds in the curves traced out by $R_\Omega(z, w)$, but I suspect that the two key principles which govern the location of C_m are already evident in the two examples above. These are as follows.

(a) C_m is constructed by the continuation of the cross block curves which are connected on the circle at infinity, and these curves will be a subset of the curves C_m^{1k}.

(b) The connection curves of Λ_1 formed by the branch Λ_1^+ are in some sense isolated from the real positive z axis by C_m.

These two constraints strongly suggest that the convergence of C_m to C_∞ is highly pathological in character involving an ever-increasing number of connections to the circle at infinity coupled with a highly intricate isolation of the connection curves of Λ_1^+ from the real axis. Evidence of such pathological behaviour has recently been observed by Wood and Turnbull (1986) in attempts to extend finite-size scaling calculations into the complex z plane. In the approach of C_m to C_∞ I therefore expect the following geometry to emerge. The cross block curves which are connected to the circle at infinity will become dense and these curves simultaneously form some isolation of C_m^1 from the real positive axis. In the limit of $m \to \infty$ the isolation curve becomes the boundary formed by C_m^1 which isolates dense connections to the circle at infinity from a line element on the positive real axis which is $0 \leq z \leq z_c$, which z_c is of course the critical point. This is to say that both the isolation curve and C_m^1 in the limit of $m \to \infty$ simultaneously converge to C_∞.

We have now arrived at the conclusion that the thermodynamic limiting locus C_∞ is in fact determined by the limit

$$C_\infty = \lim_{m \to \infty} C_m^{1+} \tag{54}$$

where C_m^{1+} is the subset of C_m^1 traced out by the branches of Λ_1 which are simultaneously maximum and equal in modulus. As is well known, the theory of critical correlations has long associated the divergence of the correlation range with the occurrence of asymptotic degeneracy of the two principal eigenvalues (z real) of the transfer matrix (Kac 1968, Fisher and Burford 1967, Thompson 1972). For z real these two eigenvalues generally belong to different blocks of (5). The structure described above would indicate that asymptotic degeneracy occurring at the critical point z_c is of two distinct algebraic types. Firstly, a subset of the branch points of Λ_1^+ will converge onto the real axis at z_c and in this the critical point is associated with a branching phenomenon and a cycle structure obtained in the limit of $m \to \infty$; this will determine the nature of the singularity of the free energy at z_c and in particular the specific heat exponent (see § 5). Secondly, the cross block curves which form the isolation curve converge onto C_∞ in the limit of $m \to \infty$ indicating that asymptotic degeneracy in modulus occurs at z_c between eigenvalues contained in different blocks, and this will determine the divergence of the correlation range and the exponent ν.

The identification of C_∞ as the limit given in (54) is the main conclusion of this section. It states that in finding the location of C_∞ it is sufficient to consider only the connection curves C_m^1 of Λ_1 which are generated by the resolvent $R(z, w)$ defined in (20) and (21) belonging to the characteristic polynomial of the block $\tau_1(z)$ in (5). It was Fisher (1965) who first observed that the original theorems of Yang and Lee (1952) and Lee and Yang (1952) which were developed in the theory of condensation in the complex activity plane could be transcribed to apply in the temperature plane to the partition function of a lattice model problem. The partition function/site of an $m \times \infty$ lattice section is a branch of Λ_1 defined by (13) and is an eigenvalue of $\tau_1(z)$. We have already remarked that $\tau_1(z)$ is a strictly positive matrix ($z > 0$) and hence the function

$$Z_{mn}^{(1)}(z) = \mathrm{Tr}[\tau_1(z)]^n \qquad (55)$$

is a polynomial in z with positive coefficients. Here $Z_{mn}^{(1)}$ is not the partition function of the $m \times n$ section but it converges to the partition function per site of the section in the limit of $n \to \infty$. Because $\tau_1(z)$ is strictly positive the Yang–Lee theorems apply not only to the whole partition function Z_{mn} but also to the block partition function $Z_{mn}^{(1)}$ alone. Thus the connection curves C_m^{1+} which satisfy condition (iii) above cannot intersect the real positive axis at any point; they are the asymptotic locus of zeros of the block partition function $Z_{mn}^{(1)}$.

The limiting partition function per site Λ is completely determined by the distribution of zeros over the limiting locus of C_m^{1+} and it is therefore sufficient to consider only C_m^{1+} in a study of partition function zeros. In addition the branch point structure of C_m^{1+} for finite m suggests a general scheme for obtaining accurate approximations to critical points. Applications of this type are the subject of II. Also, the critical point singularities are seen to be obtained as a limit of an infinite sequence of algebraic singular points and this is discussed in the following section. The factorisation (5) employed by Martin (1986) for the square lattice Potts models in the staggered ice model representation places both λ_1 and λ_2 in (42) in the same block. In such a representation C_m itself is a set of connection curves but confining a calculation to strictly finite lattices will not reveal the branch point structure of these curves which, as described above, is of both practical and theoretical significance (see II, § 3).

5. Universality and algebraic cycles

The analytic structure of $\Lambda_1^+(z)$ in the limit of $m \to \infty$ determines the thermodynamic critical point behaviour of any one-parameter model, the thermodynamic functions being defined only on the real positive z axis. The critical points and exponents are determined by the singular points of Λ_1^+ in the thermodynamic limit. In their original papers Yang and Lee (1952) and Lee and Yang (1952) identified the critical points as the intersection points of C_∞ and the real positive z axis, and showed that the critical exponents emerged in terms of the distribution of zeros on C_∞ in the neighbourhood of the critical points. The present development of the theory allows something to be said about the critical exponents without recourse to the distribution function of zeros on C_∞. It is the purpose of this section to investigate the possible values available to the critical exponents in terms of the accumulation of branch points on C_m^{1+} in the limit of $m\to\infty$. If a finite number of branch points accumulate asymptotically at the critical point then it is shown that the exponents α and ν are expected to take a specific rational form and the exponent δ to be an integer.

The analytic structure of the partition function in the thermodynamic limit emerges in the development of the function Λ_1^+ in the limit of $m\to\infty$. For all arbitrarily large and finite m, Λ_1^+ is a branch of the algebraic function Λ_1 defined by the characteristic polynomial of $\tau_1(z)$. Hence for all m, however large, Λ_1 has only a finite number of algebraic singular points. These are the branch points of Λ_1, some of which are the branch points of Λ_1^+ located on C_1^+, and Λ_1^+ in the neighbourhood of one of its branch points z_k is represented by a development of the form (27), namely

$$\Lambda_1^+(z)=\Lambda_1^+(z_k)+\sum_r a_r\varepsilon^{r/p} \tag{56}$$

where p is the cycle number and Λ_1^+ is a member of a cycle system of p branches of Λ_1. If C_∞ is formed in the manner of (54) then not only the zeros of $Z_{mn}(z)$ lie asymptotically on C_∞ but all of the branch points of Λ_1^+ close onto the same limiting locus. In the limit $m\to\infty$ the resolvent becomes an infinite polynomial and the number of branch points of Λ_1 becomes infinite. Thus we arrive at a new concept, namely the limiting distribution of the branch points of Λ_1 and in particular the subset of this distribution which belongs to Λ_1^+.

In (56) we have the immediate prospect of rational critical exponents since for all finite m they are forced to be so, arbitrarily close to the real axis. Whether the exponents survive as rational numbers under the limit of $m\to\infty$ will depend upon the limiting distribution of the branch points of Λ_1^+ on C_∞. In the two-dimensional nearest-neighbour Ising model the branch points of Λ_1^+ become dense on C_∞ (Wood 1985): in fact they are infinitely dense everywhere. This phenomenon was remarked upon by Onsager (1944) in his original paper. He observed that the critical point was a branching process 'of infinite order'. For strictly finite m, however, the branch points z_k of Λ_1^+ all have a cycle number of two, where

$$\Lambda_1^+=\Lambda_1^+(z_k)+a_1\varepsilon^{1/2}+a_2\varepsilon+\dots \tag{57}$$

in the neighbourhood of z_k, which can be arbitrarily close to z_c. The infinite density of branch points in the neighbourhood of z_c on C_∞ annihilates the algebraic character of the singularity in the limiting partition function per site. The outcome in this case is the well known logarithmic singularity, and the fact that C_∞ is also a natural boundary of the limiting partition function. In the few exact solutions available for other one-parameter lattice models the specific heat exponent is a simple fraction. The triplet Ising model of (9) and the hard hexagon gas (Baxter 1982) where $\alpha=\frac{2}{3}$ and $\frac{1}{3}$, respectively, are both examples of this. Thus if $\Lambda(z)$ is the partition function per site then to first order in $\varepsilon=z-z_c$ $\Lambda(z)$ takes the form

$$\Lambda(z)=\Lambda_c+a\varepsilon^{2-\alpha}+\dots \tag{58}$$

where z_c is an algebraic singular point of Λ. In addition there are several strong conjectures for rational exponents in two-dimensional one-parameter models such as the Potts models (for a review see Wu 1982), and many 'almost exact results' obtained from the Coulomb gas analogy (Nienhius 1984). In such cases the function $\Lambda(\varepsilon)$ would at least be an algebraic function asymptotically in the neighbourhood of $\varepsilon = 0$, whose polynomial equation becomes asymptotically of the form (58).

Of the branch points forming the connection curves C_m^{1+} as $m \to \infty$ at least one complex conjugate pair of branch points closes into the critical point z_c. In the case of the two-dimensional Ising model it is instructive to consider precisely how the algebraic singularity of (57) for finite m becomes overwhelmed in the limit of $m \to \infty$. For finite m, Λ_1^+ is given by

$$\Lambda_1^+ = (2 \sinh 2K)^{m/2}\{y_1 y_3 y_5 \ldots y_{2m-1}\} \tag{59}$$

where y_r is one of the branches of a two-valued function

$$y_r^2 - 2y_r\left(s + s^{-1} - \cos\frac{r\pi}{2m}\right) + 1 = 0 \qquad r = 1, 3, \ldots, 2m-1,\ s = \sinh 2K \tag{60}$$

(Onsager 1944). Equation (60) forms a family of quadratic equations parametrised by a single parameter r. Clearly the branch points of the functions y_r are the branch points of Λ_1^+ and lie on $|s| = 1$ (Wood 1985). In the limit of $m \to \infty$ the parametrisation r in (60) becomes the continuum and

$$\ln \Lambda = \tfrac{1}{2} \ln \sinh 2K + \frac{1}{2\pi}\int_0^\pi \ln y(\phi)\, \mathrm{d}\phi. \tag{61}$$

For m finite $y_r = 1$ at its branch point where

$$y_{r+k} \sim 1 + \mathrm{O}\left(\frac{k}{m}\right)^{1/2}. \tag{62}$$

In the limit of $m \to \infty$ this point is an accumulation point of infinitely many branch points. This pathological behaviour in the density of branch points on C_m^{1+} in the limit of $m \to \infty$ destroys the simple algebraic character of the singularity (57) which exists for all finite m.

We know that the branch points of Λ_1 become infinite in number in the limit of $m \to \infty$ and that in this limit these are distributed on the connection curves of Λ_1. We do not know how the branch points are distributed on that part of these connection curves which is given by (54). The possible outcomes are

(*a*) the branch points on C_∞ are infinitely dense everywhere,

(*b*) the branch points on C_∞ are dense with a well defined distribution everywhere and

(*c*) the branch points on C_∞ are discrete.

The critical exponents will be determined by the distribution of branch points at z_c; it seems plausible to expect that z_c will be asymptotically an algebraic singular point of $\Lambda(z)$ unless the branch points become infinitely dense at z_c in the limit of $m \to \infty$. Consider m large and finite. The connection curves C_m^{1+} form a system converging to C_∞ and initially let us suppose that the branch points of C_m^{1+} are all of cycle number 2. A schematic representation of this is given in figure 6. Let the branch points on C_m^{1+} be denoted by $z_k(m)$, $k = 1, 2, \ldots$. Each can be characterised by an integer pair $(1, j)$ $(j = 2, 3, \ldots)$ where for a given k the jth branch of Λ_1 cycles with Λ_1^+. In the limit of $m \to \infty$ let the number of branch points which converge onto z_c be

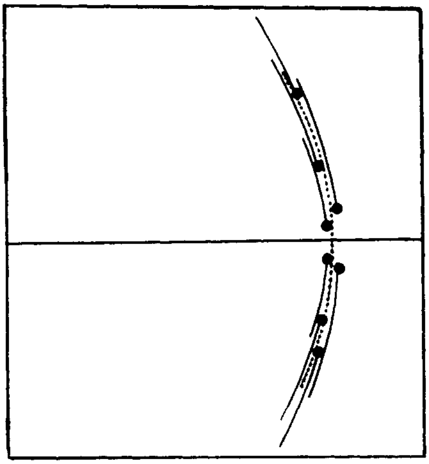


Figure 6. A schematic illustration of the convergence of the branch points of Λ_1^+ onto C_∞ represented by the dotted curve.

finite and equal to s. We can identify asymptotically branch pairs $(1, 2), (1, 3), \ldots, (1, s)$ of Λ_1 which become equal at z_c. Under such circumstances it is plausible to expect that the limiting function Λ would be algebraic (at least in the neighbourhood of z_c) where z_c is a branch point of Λ with a cycle number $p \leq s$. The denominator of the exponent in (58) (prior to expressing as an irreducible fraction) is now determined. One would generally expect this to be the upper limit of s. The numerator of the exponent depends upon the initial term in the expansion (56); if $a_1 \neq 0$ then in the limit of $m \to \infty$ z_c will be asymptotically a root of $R(z, w)$ of multiplicity $2(s-1)$, an effect induced by the meeting of $s-1$ complex conjugate pairs of connection curves at z_c. In fact the critical point is asymptotically a branch point of the branches $z_i(w)$ generated by the resolvent (see § 3 and II). The simplest value adopted by the exponent in (58) would now be

$$2-\alpha=2(s-1)/s \tag{63}$$

(for a full analysis of initial exponents in ε expansions of algebraic functions see Fuchs and Shabat (1961) and Appell and Goursat (1929)), whence we arrive at a specific heat exponent α in the form

$$\alpha=2/s \qquad s=3,\ldots. \tag{64}$$

If the scaling relation $d\nu=2-\alpha$ holds then the correlation length exponent for two-dimensional models takes the very specific form

$$\nu=(s-1)/s \qquad s=3,4,\ldots. \tag{65}$$

The two-dimensional Ising model exponents of $\alpha=0$ and $\nu=1$ are included in (64) and (65) in the limit of $s \to \infty$ corresponding to an infinite number of branch points converging onto z_c (Onsager 1944). We note that both the three-spin Ising model and the hard hexagon gas on the triangular lattice (Baxter 1982) satisfy (64) and (65) with $s=3$ and 6, respectively. These two cases also correspond to the conjectured exponents of the four- and three-state Potts models (for a review see Wu 1982). The self-avoiding walk problem in two dimensions also satisfies (65) with $s=4$ (de Gennes 1972).

The exponents in (64) and (65) assume that $a_1 \neq 0$ in (56); if the first non-zero term in (66) for a given model is a_t then in two dimensions under the same assumptions the exponents become

$$\alpha = 2(1 - t + t/s) \tag{66}$$

and

$$\nu = t(s-1)/s. \tag{67}$$

The exponents of the two-dimensional percolation problem are now also included corresponding to $t = 2$ and $s = 3$ in (66) and (67) leading to a negative specific heat exponent. We see that an infinite specific heat is only possible if $t = 1$. The three-spin Ising model on the triangular lattice (Baxter and Wu 1974, Baxter 1974) appears to provide a specific example of the above mechanism in which z_c is an algebraic singular point of $\Lambda(z)$. The model is self-dual; this is a property which allows the branch points of Λ_1^+ to accumulate on the unit circle $|z| = 1$ where $z = \sinh 2K$ (see II for further discussion). In the limit of $m \to \infty$ the partition function per site is entirely algebraic and is given by

$$\Lambda^2/6z = y \tag{68}$$

where y is the four-valued function defined by

$$z(1+z^2)(y-1)^3(1+3y) = 2(1-z)^4 y^3. \tag{69}$$

The critical point at $z = 1$ is a branch point of multiplicity 3 and cycle number 3 corresponding to $s = 3$ in (63). The first term in the ε expansion of the three branches forming a cycle at $z = 1$ is easily verified to be

$$y = 1 + \frac{1}{4^{1/3}} \varepsilon^{4/3} + \dots \tag{70}$$

confirming the numerator of $2(s-1)$ in (63). The algebraic singularities of the thermodynamic functions of this model have been very thoroughly discussed by Joyce (1975a, b). The connection curves of the four branches of y are shown in figure 7 where, as expected, the unit circle $|z| = 1$ appears.

It is of course possible that algebraic cycles larger than two appear in the branch points $z_k(m)$ of Λ_1^+ and simple equations for the exponents α and ν can again be obtained on the basis of the cycle numbers of those branch points which converge onto z_c in the limit of $m \to \infty$. Again the total number s of branches of Λ_1^+ which are asymptotically equal at $z = z_c$ can be the denominator of the exponent in (58). Here

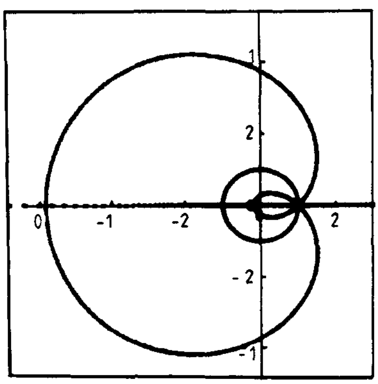


Figure 7. The connection curves of the algebraic function y in (69) in terms of which the limiting partition function per site Λ of the three-spin Ising model is defined. The connection curves of Λ_1^+ lie on the unit circle, here shown in the $z = \sinh 2K$ plane.

it would seem possible to identify a concept of 'universality' as something which is essentially algebraic in character, namely the value of s. As presently understood by this author the statement that models A and B belong to the same universality class amounts to little else than the observation that they possess the same critical exponents. Thus on the basis of numerical work it is said that the hard square gas (Baxter *et al* 1980, Wood and Goldfinch 1980) and the two-dimensional Ising model are in the same universality class. As far as this author is aware no mathematical meaning has as yet been attached to such a statement. If indeed models in statistical mechanics grouped themselves into such classes there would surely have to be a very strong mathematical mechanism promoting such groupings, and also one of some simplicity. The exponents of two-dimensional models which are simple fractions suggest strongly that this mechanism must be algebraic in origin. In the development of the limiting function $\Lambda(z)$ through an infinite sequence of algebraic functions Λ_1 with only algebraic singular points we have a basis for such a mechanism which is algebraic. It appears to the present author that the denominator of any thermodynamic exponent is much more significant than the numerator since it, or a multiple of it, is the algebraic cycle number associated with the critical point, which ultimately derives from the cycle structure of Λ. Thus in the present case of one-parameter models the numbers s and t in (63)-(67) suggest themselves as a double index for universality classes. s is the number of branches of Λ_1^+ which are asymptotically equal at z_c. For models with divergent specific heats $t=1$, and so for such cases s acts as a single indexing parameter for a universality class. Thus $s=\infty$ represents a whole class, independent of dimension, where $\alpha=0$ and $\nu=2/d$. Similarly, $s=16$ represents a whole class where $\alpha=\frac{1}{8}$ and $\nu=15/8d$. Thus if the hard square gas really is in the same universality class as the Ising model we should expect z_c for this model to be an accumulation point of infinitely many branch points in the limit of $m\to\infty$. The case $s=2$ is exceptional in that the exponent $2-\alpha$ in (58) and (63) is an integer, whence Λ is analytic at z_c, and the mean-field exponent $\alpha=0$ is recovered corresponding to a finite specific heat. In such a case in the limit of $m\to\infty$ a pair of quadratic branch points of Λ_1^+ converges onto z_c. The two-dimensional Ising model ($s=\infty$) and the mean-field model ($s=2$) mark the two extremes of this scheme for indexing universality classes, and $\alpha=\frac{2}{3}$ would be the largest positive value of α which is possible.

Finally we can attempt to extend the analysis above to models in an external field H where the critical exponent δ is defined by

$$\Lambda=\Lambda_c+a_1H^{(\delta+1)/\delta}+\ldots \qquad (z=z_c) \tag{71}$$

representing the critical behaviour to occur at $H=0$. Although some caution should be exercised in extending the present algebraic analysis to algebraic functions of more than one variable we may envisage the connection curves C_m^{1+} in the complex field plane at real values of temperature (see II, § 8). In the limit of $m\to\infty$ branch points will accumulate onto a limiting locus and δ in (71) can therefore be a cycle number of the branch points accumulating at $H=0$, and hence an integer. For the two-dimensional Ising, Potts and hard hexagon models δ is either known exactly or conjectured to be an integer.

6. Summary

This paper has considered some of the algebraic features which are common to all the one-parameter lattice models of statistical mechanics. An understanding of the way in which the limiting locus of partition function zeros C_∞ emerges as a limiting process defined on a sequence of $m\times\infty$ lattices has been attempted. The starting point is the observation that the eigenvalues of finite transfer matrices T_Ω are all algebraic functions defined by a factorisation of the characteristic equation. The partition function per site Λ_1^+ at real temperatures is a branch of an algebraic function Λ_1 defined by the characteristic equation of a strictly positive matrix τ_1 which can be obtained in a block diagonalisation of T_Ω. For one-parameter models such a factorisation is always possible.

In § 3 we have defined a cut L in the complex z plane, where at all points in the cut plane a many-valued algebraic function $\Lambda(z)$ defined by a irreducible polynomial equation in Λ and z has function elements which are distinct in modulus. The cut L is a set of connections on the branch points of Λ which become linked via a system of algebraic curves which themselves are the branches of an algebraic function defined by the polynomial resolvent of (19) and (20). These curves have been called the connection curves of the algebraic function Λ; they exhibit a highly structured geometry in terms of their intersection points which correspond to the occurrence of branch points in the algebraic functions generated by the resolvent equation. Everywhere on a connection curve there exists a real quadratic factor in the original defining polynomial of Λ.

In § 4 we have considered C_∞ as the limit of the sequences C_m and C_m^{1+} which are traced out by the algebraic curves generated by the resolvent (19) when constructed from the characteristic equations of the whole transfer matrix T_Ω and the block matrix respectively. These algebraic curves are points in the z plane where the eigenvalues of T_Ω of τ_1 are simultaneously equal and maximum in modulus, but only those obtained from τ_1 form a cut of the type L discussed in § 3. It is postulated that both C_m and C_m^{1+} converge onto C_∞. Since only C_m^{1+} is linked to the branch point structure of Λ_1 and since the thermodynamic critical point is determined by the convergence of branch points onto z_c it is of greater significance and interest and is also easier to compute.

Possible implications for the value of the critical exponents α, ν and δ are developed in § 5. For all finite but arbitrarily large m Λ_1^+ is an algebraic function and consequently has only algebraic singular points. In the neighbourhood of any such singular point z_k Λ_1^+ has a unique expansion in fractional powers of $\varepsilon=z-z_k$. In the limit of large m Λ_1^+ has an algebraic singular point in the complex plane which can be made arbitrarily close to the real critical point z_c. Thus we envisage that the singularity at z_c in the limiting partition function per site Λ is approached through a sequence of the algebraic singular points of Λ_1^+ which themselves converge onto C_∞. It is suggested that the nature of the singularity in Λ at z_c up to first order in ε will be determined by the number of the branch points of Λ_1^+ which in the limit of $m\to\infty$ converge onto the critical point z_c, and that this number s is a natural index for a universality class. It is assumed that for s finite the singularity in Λ at z_c will be algebraic and that, subject to simple assumptions on the ε expansion about z_c, the exponent ν may commonly be expected to take the specific rational form (65) for many two-dimensional one-parameter lattice models, and the exponent δ to be an integer.

Acknowledgment

I am grateful to Professor David Burgess for many helpful discussions.

References

Ahlfors L V 1953 *Complex Analysis* (New York: McGraw-Hill)
Appell P and Goursat E 1929 *Théorie des Fonctions Algébriques et de leurs Intégrales* (Paris: Gauthier-Villars)
Baxter R J 1974 *Aust. J. Phys.* **27** 369
—— 1982 *Exactly Solved Models in Statistical Mechanics* (New York: Academic)
Baxter R J, Enting I G and Tsang S K 1980 *J. Stat. Phys.* **22** 465
Baxter R J and Wu F Y 1974 *Aust. J. Phys.* **27** 357
de Gennes P G 1972 *Phys. Lett.* **38A** 339
Derrida B, De Seze L and Itzykson C 1983 *J. Stat. Phys.* **33** 559
Fisher M E 1965 *Lectures in Theoretical Physics* vol 7c (Boulder, CO: University of Colorado Press)
Fisher M E and Burford R J 1967 *Phys. Rev.* **156** 583
Fuchs B A and Shabat B V 1961 *Functions of a Complex Variable* vol 2 (Oxford: Pergamon)
Goursat E 1916 *Functions of a Complex Variable* (Aylesbury: Ginn)
Hardy G H and Wright E M 1960 *An Introduction to the Theory of Numbers* (Oxford: Clarendon)
Joyce G S 1975a *Proc. R. Soc.* A **343** 45
—— 1975b *Proc. R. Soc.* A **345** 277
Kac M 1968 *Mathematical Mechanisms of Phase Transitions, Brandeis Lectures 1966* (New York: Gordon and Breach) p 243
Lee T D and Yang C N 1952 *Phys. Rev.* **87** 410
Littlewood D E 1958 *A University Algebra* (London: Heinemann)
Martin P 1986 *J. Phys. A: Math. Gen.* **19** 3267
Nienhuis B 1984 *J. Stat. Phys.* **34** 731
Onsager L 1944 *Phys. Rev.* **65** 117
Ree F H and Chestnut D A 1966 *J. Chem. Phys.* **45** 3983
Runnels L K and Combs L L 1966 *J. Chem. Phys.* **45** 2482
Thompson C J 1972 *Mathematical Statistical Mechanics* (London: MacMillan)
Turnbull H W 1952 *Theory of Equations* (Edinburgh: Oliver and Boyd)
Wood D W 1985 *J. Phys. A: Math. Gen.* **18** L481
Wood D W and Goldfinch M C 1980 *J. Phys. A: Math. Gen.* **13** 2781
Wood D W and Turnbull R W 1986 *J. Phys. A: Math. Gen.* **19** 2611
Wood D W, Turnbull R W and Ball J K 1987 *J. Phys. A: Math. Gen.* **20** 3495
Wu F Y 1982 *Rev. Mod. Phys.* **54** 235
Yang C N and Lee T D 1952 *Phys. Rev.* **87** 404, 410