Discrete Mathematics 259 (2002) 37–57

DISCRETE MATHEMATICS

www.elsevier.com/locate/disc

Equimodular curves

Norman Biggs

Centre for Discrete and Applicable Mathematics, London School of Economics, Houghton Street, London WC2A 2AE, UK

Received 22 March 2001; received in revised form 2 January 2002; accepted 14 January 2002

Abstract

This paper is motivated by a problem that arises in the study of partition functions of Potts models, including as a special case chromatic polynomials. When the underlying graphs have the form of 'bracelets', the chromatic polynomials can be expressed in terms of the eigenvalues of a matrix. In this situation a theorem of Beraha, Kahane and Weiss asserts that the zeros of the polynomials approach the curves on which the matrix has two eigenvalues with equal modulus. It is shown here that (in general) these 'equimodular' curves comprise a number of segments, the end-points of which are the roots (possibly coincident) of a polynomial equation. The equation represents the vanishing of a discriminant, and the segments are in bijective correspondence with the double roots of another polynomial equation, which is significantly simpler than the discriminant equation. Singularities of the segments can occur, corresponding to the vanishing of a Jacobian. In addition, it is proved by algebraic means that the equimodular curves for a reducible matrix are closed curves. The question of dominance is investigated, and a method of constructing the dominant equimodular curves for a reducible matrix is suggested. These results are illustrated by explicit calculations in a specific case.

Keywords: Chromatic polynomial; Zeros

1. Introduction

This paper is motivated by a problem that arises in the study of partition functions of Potts models, including as a special case the chromatic polynomial. The critical behaviour of these models leads to the study of the complex zeros of the partition function, the classic result in this field being the Lee–Yang theorem [8,9].

E-mail address: n.biggs@maths.lse.ac.uk (N. Biggs).

PII: S0012-365X(02)00444-2

In 1972 it was observed [3] that the complex zeros of the chromatic polynomials of certain graphs ('bracelets') exhibit interesting behaviour, although the reason for it was not understood at that time. Subsequently, a theorem of Beraha et al. [1] provided a general explanation for such behaviour. (Specific examples are given in [12].) The theorem asserts that, as $n \to \infty$, the zeros of certain sequences of polynomials $f_n(z)$ approach the curves on which a matrix $A(z)$ has two eigenvalues with equal modulus.

In the theoretical physics literature, the Beraha–Kahane–Weiss theorem has been exploited in papers by Chang et al. [5,6,14,15], and by Salas and Sokal [13]. It is clear that it requires intensive computational resources, and consequently some theoretical work on it is desirable. That is the subject of the present paper.

The main topic is a function $v_A(t,z)$, a polynomial in the real variable t and the complex variable z. The points z where $A(z)$ has two eigenvalues with equal modulus are those for which $v_A(t,z)=0$ for some t in the real interval $[0,4]$ (Theorem 1). It will be shown that the resulting 'equimodular curves' comprise a number of segments, the end-points of which correspond to the vanishing of a discriminant. (The author is grateful to a referee for pointing out that this situation was discussed in the 1980s, in papers by Wood [17] and Martin [10].) Here it is proved that the segments are in bijective correspondence with the double roots of another polynomial equation, which is significantly simpler than the discriminant equation (Theorem 2). The segments may contain singularities, one common type being the 'real-crossing' singularity, which is described in detail.

If the matrix $A(z)$ is reducible, its equimodular curves can be constructed by looking at its constituents. It is shown that the curves arising from the interaction between constituents are closed curves (Theorem 3).

Finally we discuss the question of identifying the equimodular curves that are dominant, in a certain sense. A feature of this discussion is the existence of 'triple points', and in the reducible case these points play a particularly important part (Theorem 4).

The algebraic and computational techniques employed here are part of the current revival of interest in the techniques of classical algebraic geometry, in particular the use of resultants [7]. The computations in Section 9 were done with the aid of Maple, version 6, and I am grateful to Philipp Reinfeld for help with them.

2. The polynomial criterion

Given a function A, from the complex field $\mathbb{C}$ to the ring of $m \times m$ matrices with complex entries, we define the *equimodular set* $E(A)$ as follows. $E(A)$ is the set of $\zeta \in \mathbb{C}$ for which there is a neighbourhood N of ζ and two distinct complex-valued functions λ_1, λ_2 defined on N satisfying

$$|\lambda_1(\zeta)| = |\lambda_2(\zeta)| \quad \text{and} \quad \det(\lambda_i(z)I - A(z)) = 0 \quad (i = 1, 2,\ z \in N).$$

In other words, $E(A)$ is the set of points ζ where $A(\zeta)$ has two eigenvalues of equal modulus. The definition is formulated so that these eigenvalues may in fact be equal at ζ, but they must not be identically equal. The following theorem implies that, under certain conditions, $E(A)$ is a set of curves in the complex plane. These are the 'equimodular curves' of the title.

Theorem 1. *Let A be a function from the complex field $\mathbb{C}$ to the ring of $m \times m$ matrices with complex entries, such that each entry of the matrix $A(z)$ is a polynomial function of z with integer coefficients. Then there is a function $v_A : \mathbb{R} \times \mathbb{C} \to \mathbb{C}$, polynomial in both variables and with integer coefficients, such that the equimodular set for A is given by*

$$E(A) = \{z \in \mathbb{C} \mid v_A(t,z) = 0, \text{ for some } t \in [0,4]\}.$$

The proof of the theorem, including construction of the polynomial v_A, will occupy the next two sections. The key idea is the observation that if $|\lambda_1| = |\lambda_2|$, then there is a complex number s with $|s| = 1$ such that $\lambda_1 = s\lambda_2$. So if λ_1 and λ_2 are roots of the characteristic equation $a(\lambda) = \det(\lambda I - A(z)) = 0$, then it follows that λ_2 is a common root of $a(\lambda) = 0$ and $a_s(\lambda) = 0$, where $a_s(\lambda) = a(s\lambda)$.

Let $a_i(z)$ be the coefficient of λ^{m-i} in $a(\lambda)$, so that

$$a(\lambda) = \det(\lambda I - A(z)) = \lambda^m + a_1(z)\lambda^{m-1} + a_2(z)\lambda^{m-2} + \cdots + a_m(z),$$

$$a_s(\lambda) = s^m\lambda^m + s^{m-1}a_1(z)\lambda^{m-1} + s^{m-2}a_2(z)\lambda^{m-2} + \cdots + a_m(z).$$

The coefficients $a_i(z)$ are the sums of principal minors of $A(z)$ and so, under the conditions of the theorem, they are polynomials with integer coefficients. The properties of $E(A)$ depend on these coefficients, rather than the individual entries of A, and so the discussion will focus on the functions $a_1, a_2, \ldots, a_m$.

In fact, it is convenient to begin by regarding the a_i as indeterminates, making the obvious convention that $a_0 = 1$. It is a classical result that a necessary and sufficient condition for the polynomials $a_s(\lambda)$ and $a(\lambda)$ to have a non-constant common factor is that the *resultant* det R vanishes, where $R = R(a_s, a)$ is the following $2m \times 2m$ matrix:

$$\begin{pmatrix}
s^m & s^{m-1}a_1 & \ldots & sa_{m-1} & a_m & 0 & \ldots & 0 & 0 \\
0 & s^m & \ldots & s^2a_{m-2} & sa_{m-1} & a_m & \ldots & 0 & 0 \\
0 & 0 & . & s^3a_{m-3} & s^2a_{m-2} & sa_{m-1} & \ldots & 0 & 0 \\
. & . & . & . & . & . & . & . & . \\
. & . & . & . & . & . & . & . & . \\
0 & 0 & \ldots & s^m & s^{m-1}a_1 & s^{m-2}a_2 & \ldots & sa_{m-1} & a_m \\
1 & a_1 & \ldots & a_{m-1} & a_m & 0 & \ldots & 0 & 0 \\
0 & 1 & \ldots & a_{m-2} & a_{m-1} & a_m & \ldots & 0 & 0 \\
0 & 0 & \ldots & a_{m-3} & a_{m-2} & a_{m-1} & \ldots & 0 & 0 \\
. & . & . & . & . & . & . & . & . \\
. & . & . & . & . & . & . & . & . \\
0 & 0 & \ldots & 1 & a_1 & a_2 & \ldots & a_{m-1} & a_m
\end{pmatrix}.$$

By definition, $\det R$ is the sum over all permutations π of $\{1,2,\ldots,2m\}$ of terms

$$\mathrm{sign}(\pi)\ r_{1,\pi(1)}r_{2,\pi(2)}\cdots r_{2m,\pi(2m)}.$$

For $1\leqslant i\leqslant m$ the non-zero entries of R in rows i and $i+m$ are in columns $i,i+1,\ldots i+m$. Thus the non-zero terms in $\det R$ arise only from permutations that satisfy the condition

$$\{\pi(i),\pi(i+m)\}\subset\{i,i+1,\ldots,i+m\}.$$

We shall denote the set of such permutations by Π_m.

We can describe non-zero entries of R as follows:

$$r_{ij}=\begin{cases} s^{j-i+m}a_{j-i} & \text{if } 1\leqslant i\leqslant m;\\ a_{j-i+m} & \text{if } m+1\leqslant i\leqslant 2m.\end{cases}$$

It follows that for each $\pi\in\Pi_m$, there are non-negative integers $n_1(\pi)$, $n_2(\pi)$, …, $n_m(\pi)$ and a non-negative integer $e(\pi)$ such that

$$\det R=\sum_{\pi\in\Pi_m}\mathrm{sign}(\pi)a_1^{n_1(\pi)}a_2^{n_2(\pi)}\cdots a_m^{n_m(\pi)}s^{e(\pi)}.$$

Thus $\det R(a_s,a)$ may be regarded as a polynomial in s whose coefficients are integral linear combinations of monomials in the indeterminates $a_1,a_2,\ldots,a_m$. We shall refer to this polynomial as the *generic* polynomial $\rho_m(s)$.

In order to describe the coefficients of $\rho_m(s)$, let $[n]=(n_1,n_2,\ldots,n_m)$, and

$$a[n]=a_1^{n_1}a_2^{n_2}\ldots a_m^{n_m}.$$

We shall identify $[n]$ with the partition of the integer $n_1+2n_2+\cdots+mn_m$ in which n_i parts are equal to i $(1\leqslant i\leqslant m)$.

Lemma 1. *If the monomial $a[n]$ occurs in $\rho_m(s)$ then $[n]$ is a partition of m^2 such that no part is greater than m, and not more than m parts are equal.*

Proof. Given $\pi\in\Pi_m$, define functions

$$t_\pi,b_\pi:\{1,2,\ldots,m\}\to\{0,1,\ldots,m\}$$

by the rules

$$t_\pi(i)=\pi(i)-i,\quad b_\pi(i)=\pi(i+m)-i.$$

Consider the term in $\rho_m(s)$ arising from a given $\pi\in\Pi_m$. For $1\leqslant i\leqslant m$ the element $r_{i,\pi(i)}$ involves a_k if and only if $k=\pi(i)-i$, and the element $r_{i+m,\pi(i+m)}$ involves a_k if and only if $k=\pi(i+m)-i$. Thus

$$n_k(\pi)=\#t_\pi^{-1}(k)+\#b_\pi^{-1}(k).$$

It follows that

$$\sum_{k=0}^{m}kn_k(\pi)=\sum_{i=1}^{m}(t_\pi(i)+b_\pi(i)).$$

Now

$$\sum_{i=1}^{m}(t_\pi(i)+b_\pi(i)) = \sum_{i=1}^{m}(\pi(i)-i) + \sum_{i=1}^{m}(\pi(i+m)-i)$$
$$= \sum_{i=1}^{2m}\pi(i) - 2\sum_{i=1}^{m} i.$$

Since π is a permutation, this is equal to

$$\sum_{i=1}^{2m} i - 2\sum_{i=1}^{m} i = (2m^2+m) - 2(m(m+1)/2) = m^2.$$

Hence $[n](\pi) = (n_1(\pi), n_2(\pi), \ldots, n_m(\pi))$ is a partition of m^2, satisfying the stated conditions. □

If $[n](\pi) = [n]$ we shall say that a permutation $\pi \in \Pi_m$ *induces* the partition $[n]$. We can collect the terms in $\rho_m(s)$ as follows:

$$\rho_m(s) = \sum_{[n]} a[n]\rho_m([n], s),$$

where the sum is over all partitions of the kind specified in Lemma 1, and

$$\rho_m([n], s) = \sum_{\pi \text{ induces } [n]} \text{sign}(\pi) s^{e(\pi)}.$$

Lemma 2. *If $\alpha_i[n]$ is the coefficient of s^{m^2-i} in $\rho_m([n], s)$ then*

$$\alpha_{m^2-i}[n] = (-1)^m \alpha_i[n].$$

Proof. Given $\pi \in \Pi_m$, define π^* as follows:

$$\pi^*(i) = \begin{cases} \pi(i+m) & \text{if } 1 \leqslant i \leqslant m, \\ \pi(i-m) & \text{if } m+1 \leqslant i \leqslant 2m. \end{cases}$$

Then π^* is in Π_m and $\pi^* \neq \pi$. We shall show that

$$\text{sign}(\pi^*) = (-1)^m \,\text{sign}(\pi), \quad [n](\pi^*) = [n](\pi), \quad e(\pi^*) = m^2 - e(\pi).$$

Let τ_i $(1 \leqslant i \leqslant m)$ denote the transposition that switches $\pi(i)$ and $\pi(i+m)$. Then $\pi^* = \tau_1\tau_2\ldots\tau_m\pi$, so $\text{sign}(\pi) = (-1)^m \text{sign}(\pi^*)$.

It follows from the definitions that $t_{\pi^*} = b_\pi$ and $b_{\pi^*} = t_\pi$. Hence $n_k(\pi^*) = n_k(\pi)$ for $1 \leqslant k \leqslant m$.

Finally, s^k occurs in a term $r_{i,\pi(i)}$ if and only if $1 \leqslant i \leqslant m$ and $k = m - (\pi(i) - i)$. Hence, using the formula obtained in Lemma 1,

$$e(\pi) = \sum_{i=1}^{m} (m - t_\pi(i)) = m^2 - \sum_{i=1}^{m} t_\pi(i) = \sum_{i=1}^{m} b_\pi(i),$$

and

$$e(\pi^*) = \sum_{i=1}^{m} (m - t_{\pi^*}(i)) = m^2 - \sum_{i=1}^{m} b_\pi(i).$$

Hence $e(\pi) + e(\pi^*) = m^2$.

Thus π and π^* induce the same partition, and if π contributes to the coefficient of s^i then π^* contributes to the coefficient of s^{m^2-i}, with the same sign if m is even but opposite sign if m is odd. □

The result can be expressed by the equation

$$s^{m^2} \rho_m([n], s^{-1}) = (-1)^m \rho_m([n], s).$$

Equivalently, we can write $\rho_m([n], s)$, when m is odd, in the form

$$\alpha_0[n](s^{m^2} - 1) + \alpha_1[n](s^{m^2-1} - s) + \cdots + \alpha_M[n](s^{M+1} - s^M),$$

where $M = (m^2 - 1)/2$. (There is a slightly different expression if m is even.) In the next section we shall show that the coefficients satisfy certain relations, and discuss the case $m = 3$ in detail.

3. Simplification of the condition

Lemma 3. *The generic polynomial* $\rho_m(s) = \det R(a_s, a)$ *factorizes as*

$$\rho_m(s) = a_m(s-1)^m \delta_m(s),$$

where $\delta_m(s)$ *is a reciprocal polynomial of degree* $m(m-1)$.

Proof. The factors follow immediately from the explicit form of $R(a_s, a)$. Subtract row $m+i$ from row i, for each i in the range $1 \leqslant i \leqslant m$. The resulting non-zero entries are, in row i and column j $(i \leqslant j \leqslant m+i-1)$,

$$(s^{m-j+i} - 1)a_{j-i} = (s-1)(1 + s + \cdots + s^{m-j+i-1})a_{j-i}.$$

The other entries in row i are zero. In particular, the new entry in row m and column $2m$ is zero, so there is only one non-zero entry in that column, which is a_m in the last row. Expanding in terms of the last column and removing the factor $(s-1)$ from each of the terms in rows 1 to m, we get

$$\rho_m(s) = a_m(s-1)^m \delta_m(s),$$

where $\delta_m(s)$ is a determinant of size $2m-1$.

In the previous section we showed that $s^{m^2}\rho_m(s^{-1}) = (-1)^m \rho_m(s)$. Thus

$$s^{m^2} a_m(s^{-1}-1)^m \delta_m(s^{-1}) = (-1)^m a_m(s-1)^m \delta_m(s).$$

This implies that

$$s^{m(m-1)}\delta_m(s^{-1}) = \delta_m(s),$$

which means that $\delta_m(s)$ is a reciprocal polynomial. □

The following corollary is no surprise, because when $s=1$ the condition $\lambda_1 = s\lambda_2$ means that the equation $a(\lambda)=0$ has a double root.

Corollary. *$\delta_m(1)$ is an integer multiple of the discriminant of $a(\lambda)$.*

Proof. In each of the rows 1 to m of $\delta(1)$ the non-zero entries are as follows.

$$m \quad (m-1)a_1 \quad (m-2)a_2 \quad \ldots \quad 2a_{m-2} \quad a_{m-1}.$$

These are the coefficients of $a'(\lambda)$, so $\delta(1)$ is the resultant of $a'(\lambda)$ and $a(\lambda)$, which is a multiple of the discriminant of $a(\lambda)$. □

Since $\delta_m(s)$ is a reciprocal polynomial of degree $m(m-1)$ can write

$$\delta_m(s) = s^{m(m-1)/2}\left(\beta_0 + \sum_{i=1}^{m(m-1)/2} \beta_i(s^i + s^{-i})\right).$$

The common factor a_m has been removed, so the coefficients β_i are integral linear combinations $\sum \beta_i[n']a[n']$, where the sum is now over partitions $[n']$ of $m(m-1)$ with the appropriate properties. The integers $\beta_i[n']$ can be expressed in terms of the integers $\alpha_i[n]$ by using the formula $\rho_m(s) = a_m(s-1)^m\delta_m(s)$. This formula also implies that the $\alpha_i[n]$ satisfy certain relations. (See the discussion of the case $m=3$ below.)

Further simplification occurs when the variable s is replaced by $t = s + s^{-1} + 2$. The condition $|s|=1$ implies that $s = \exp(\mathrm{i}\theta)$, where θ is real, and so $t = 4\cos^2(\theta/2)$. Hence t is real and lies in the range $0 \leqslant t \leqslant 4$. The identity

$$s^k + s^{-k} = (s+s^{-1})^k - \sum_{i=1}^{\lfloor k/2 \rfloor} \binom{k}{i} (s^{k-2i} + s^{-k+2i})$$

implies that for each $k>0$, $s^k + s^{-k}$ can be expressed as a polynomial $\phi_k(t)$ with integer coefficients. In fact, defining $\phi_0(t)=1$, we have the recursion

$$\phi_k(t) \;=\; (t-2)^k - \sum_{i=1}^{\lfloor k/2 \rfloor} \binom{k}{i} \phi_{k-2i}(t).$$

For example,

$$\phi_1(t) = t - 2, \quad \phi_2(t) = t^2 - 4t + 2, \quad \phi_3(t) = t^3 - 6t^2 + 9t - 2.$$

It follows that $\delta_m(s) = s^{m(m-1)/2} r_m(t)$, where $t = s + s^{-1} + 2$ and

$$r_m(t) = \sum_{i=0}^{m(m-1)/2} b_i \phi_i(t)$$

is a polynomial of degree $m(m-1)/2$.

We can now complete the proof of Theorem 1. Given the $m \times m$ matrix-valued function $A(z)$, the functions $a_1(z), a_2(z), \ldots, a_m(z)$ are the coefficients of its characteristic polynomial. Replacing each monomial $a[n]$ that occurs in $r_m(t)$ by the function

$$a[n](z) = a_1(z)^{n_1} a_2(z)^{n_2} \ldots a_m(z)^{n_m},$$

we have the polynomial $v_A(t,z)$. As a polynomial in t it has degree $m(m-1)/2$, and as a polynomial in z its degree depends upon the degrees of the polynomials $a_i(z)$.

We have shown how to construct $v_A(t,z)$ from the generic polynomial $r_m(t)$, which can be constructed (in theory) by combinatorial means.

Example. When $m = 3$ there are five partitions of $m^2 = 9$ in which no part exceeds 3 and no part is repeated more than three times. The corresponding monomials, that contribute to $\rho_3(s)$, are

$$a_3^3, \quad a_1 a_2 a_3^2, \quad a_1^3 a_3^2, \quad a_2^3 a_3, \quad a_1^2 a_2^2 a_3.$$

We shall determine the contribution of the monomial $a_1 a_2 a_3^2$ to $\rho_3(s)$, and hence the contribution of $a_1 a_2 a_3$ to $r_3(t)$.

According to the theory developed in Section 2, we have to enumerate the permutations π that induce the partition $(1,1,2)$. Consider the associated functions $t = t_\pi$ and $b = b_\pi$. These functions must take the values $1,2,3$ (respectively) $1,1,2$ times, and hence the value 0 twice. Thus

$$\{t(1), t(2), t(3), b(1), b(2), b(3)\} = \{0,0,1,2,3,3\}.$$

Because of the symmetry between π and π^* it is only necessary to consider the cases where $t(1) + t(2) + t(3)$ is one of the numbers $0,1,2,3,4$. In fact 0 cannot occur and there are just five possibilities for $\{t(1), t(2), t(3)\}$:

$$\{0,0,1\}, \quad \{0,0,2\}, \quad \{0,0,3\}, \quad \{0,1,2\}, \quad \{0,1,3\}.$$

The corresponding exponents $e = 9 - t(1) + t(2) + t(3)$ are $8,7,6,6,5$.

It remains to determine the order in which these values can be assigned to $t(1)$, $t(2)$, $t(3)$, and the complementary set of values to $b(1)$, $b(2)$, $b(3)$, in such a way that t

and b determine a permutation π. It suffices to work out the possibilities for $e=8,7,5$, which are given in the following table.

$t(1)$	$t(2)$	$t(3)$	$b(1)$	$b(2)$	$b(3)$	e	π	sign
0	0	1	2	2	3	8	(34)	−
0	0	2	3	1	3	7	(35)	−
0	2	0	1	3	3	7	(24)	−
0	1	3	3	0	2	5	(2365)	−
0	3	1	2	0	3	5	(25)(34)	+
1	3	0	0	2	3	5	(1254)	−

The possibilities for $e=6$ are not needed because (as indicated above) the coefficients $\alpha_i[n]$ $(0\leqslant i\leqslant 4)$ of $\rho_3([n],s)$ are not independent. In fact the existence of the factor $(s-1)^3$ means that

$$9\alpha_0[n]+7\alpha_1[n]+5\alpha_2[n]+3\alpha_3[n]+\alpha_4[n]=0.$$

Our calculations have shown that, when $[n]=(112)$, $\alpha_0[n]=0$, $\alpha_1[n]=-1$, $\alpha_2[n]=-2$, and $\alpha_4[n]=-1$. Hence $\alpha_3[n]=6$ and the contribution of the monomial $a[n]=a_1a_2a_3^2$ to $\rho_3(s)$ is

$$a_1a_2a_3^2(-(s^8-s)-2(s^7-s^2)+6(s^6-s^3)-(s^5-s^4)).$$

Removing the factors $a_3(s-1)^3$ we get the contribution of $a_1a_2a_3$ to $\delta_3(s)$:

$$a_1a_2a_3(-(s^5+s)-5(s^4+s^2)-6s^3).$$

Dividing by s^3 and making the substitution $t=s+s^{-1}+2$ we get the contribution of $a_1a_2a_3$ to $r_3(t)$:

$$a_1a_2a_3(-\phi_2(t)-5\phi_1(t)-6)=-a_1a_2a_3(t-1)(t+2).$$

Carrying out the same process for the other relevant monomials leads to the formula:

$$r_3(t)=(t-1)^3a_3^2-(t-1)(t+2)a_1a_2a_3+ta_2^3+ta_1^3a_3-a_1^2a_2^2.$$

Of course, the combinatorial method is only useful for theoretical purposes. In practice, modern computer–algebra systems, such as Maple, will evaluate $\rho_3(s)$ directly as the determinant of 6×6 matrix, and perform the subsequent algebraic reductions, instantaneously.

Putting $t=4$ gives

$$r_3(4)=27a_3^2-18a_1a_2a_3+4a_2^3+4a_1^3a_3-a_1^2a_2^2$$

which is (apart from an integer factor) the discriminant of the cubic polynomial $a(\lambda)=\lambda^3+a_1\lambda^2+a_2\lambda+a_3$. Since $t=4$ corresponds to $s=1$, this is consistent with the Corollary to Lemma 3. Putting $t=0$ gives $r_3(0)=-(a_3-a_1a_2)^2$, which is also part of a general pattern, to be explained in the next section.

4. The square property

The values $t=0$ and 4 are obviously special. The value $t=4$ corresponds to $s=1$, when the equation $a(\lambda)=0$ has roots $\lambda_1=\lambda_2$, and it follows that $r_m(4)$ is a multiple of the discriminant. The value $t=0$ corresponds to $s=-1$, and in this case the equation $a(\lambda)=0$ has roots λ_1,λ_2 with $\lambda_1=-\lambda_2$. Trivially, it is also true that $\lambda_2=-\lambda_1$, so we should expect the resultant of $a(\lambda)$ and $a_{-1}(\lambda)$ to have double roots. In this section we shall show that $\pm r_m(0)$ is indeed a perfect square.

Let

$$f(\lambda)=\frac{1}{2}(a(\lambda)+a(-\lambda)), \quad g(\lambda)=\frac{1}{2\lambda}(a(\lambda)-a(-\lambda)).$$

Then if m is even

$$f(\lambda)=\lambda^m+a_2\lambda^{m-2}+\cdots+a_{m-2}\lambda^2+a_m,$$

$$g(\lambda)=a_1\lambda^{m-2}+a_3\lambda^{m-4}+\cdots+a_{m-3}\lambda^2+a_{m-1},$$

and if m is odd

$$f(\lambda)=a_1\lambda^{m-1}+a_3\lambda^{m-3}+\cdots+a_{m-1}\lambda^2+a_m,$$

$$g(\lambda)=\lambda^{m-1}+\cdots+a_{m-3}\lambda^2+a_{m-1}.$$

In both cases we can consider f and g as polynomials in $\mu=\lambda^2$. They have a common root $\mu_1=\lambda_1^2$, and so the resultant of f and g vanishes. The resultant is the determinant of an $(m-1)\times(m-1)$ matrix S_m, which, in the even case, is

$$S_m=\begin{pmatrix} 1 & a_2 & a_4 & \dots & a_m & 0 & \dots & 0 \\ 0 & 1 & a_2 & \dots & a_{m-2} & a_m & \dots & 0 \\ . & . & . & \dots & . & . & \dots & . \\ 0 & 0 & 0 & \dots & a_4 & a_6 & \dots & a_m \\ a_1 & a_3 & a_5 & \dots & 0 & 0 & \dots & 0 \\ 0 & a_1 & a_3 & \dots & a_{m-1} & 0 & \dots & 0 \\ . & . & . & \dots & . & . & \dots & . \\ 0 & 0 & 0 & \dots & a_3 & a_5 & \dots & a_{m-1} \end{pmatrix}.$$

There is a similar form when m is odd. For example, when $m=3,4,5$, the matrices are

$$S_3=\begin{pmatrix} 1 & a_2 \\ a_1 & a_3 \end{pmatrix}, \quad S_4=\begin{pmatrix} 1 & a_2 & a_4 \\ a_1 & a_3 & 0 \\ 0 & a_1 & a_3 \end{pmatrix}, \quad S_5=\begin{pmatrix} 1 & a_2 & a_4 & 0 \\ 0 & 1 & a_2 & a_4 \\ a_1 & a_3 & a_5 & 0 \\ 0 & a_1 & a_3 & a_5 \end{pmatrix}.$$

The argument given above shows that a necessary and sufficient condition for $a_{-1}(\lambda)$ and $a(\lambda)$ to have a common root is that $\det S_m = 0$. This condition is related to our general framework by the following result.

Lemma 4. *If S_m is the $(m-1)\times(m-1)$ matrix described above, and $r_m(t)$ is the polynomial of degree $m(m-1)/2$ defined in Section 3, then*

$$r_m(0) = (-1)^{m(m+1)/2}(\det S_m)^2.$$

Proof. Let $R = R(a_{-1}, a)$ and

$$X = \frac{1}{2}\begin{pmatrix} I & I \\ -I & I \end{pmatrix},$$

where the submatrices are all of size $m \times m$. Then XR has one non-zero entry (1) in the first column, and one non-zero entry (a_m) in the last column. Expanding $\det XR$ in terms of these two columns gives

$$\det XR = a_m \det(S_m \otimes I_2),$$

where $\otimes$ denotes the Knonecker product, and I_2 is the 2×2 identity matrix. Since $\det X = 2^{-m}$ and $\det(S_m \otimes I_2) = (\det S_m)^2$, we have

$$\det R = 2^m a_m (\det S_m)^2.$$

Putting $s = -1$ in Lemma 3, we have $\det R = a_m(-2)^m \delta_m(-1)$, and from the definition of r_m, $\delta_m(-1) = (-1)^{m(m-1)/2} r_m(0)$. The result follows. □

Theorem 2. *With the conditions as stated in Theorem 1, there is a polynomial f_A with integer coefficients such that*

$$v_A(0,z) = f_A(z)^2. \qquad \square$$

5. Segments and singularities

In general, the polynomial $v_A(t,z)$ can be expressed in the form $w_A(z)u_A(t,z)$, where $w_A(z)$ is a polynomial independent of t and

$$u_A(t,z) = p_0(t)z^h + p_1(t)z^{h-1} + \cdots + p_h(t),$$

the coefficients $p_i(t)$ $(0 \leqslant i \leqslant h)$ being polynomials in t with integer coefficients.

If ζ is such that $w_A(\zeta) = 0$, then $v_A(t,\zeta) = 0$ for all $t \in [0,4]$, and the point ζ is in $E(A)$. We shall refer to such points as *degenerate arcs*.

Provided that $p_0(t_0) \neq 0$, the equation $u_A(t_0, z) = 0$ is a polynomial equation of degree h in z. Let z_0 be any one of its h roots. By the implicit function theorem, if the Jacobian of the mapping $z \mapsto u_A(t_0, z)$ is not zero at z_0, then there is a neighbourhood N of t_0 and a continuously differentiable function $z(t)$ defined on N, such that $z(t_0) = z_0$ and $u_A(t, z(t))$ is identically zero for $t \in N$. This means that, apart from the degenerate arcs,

$E(A)$ is a union of h images of the interval $[0,4]$. These images are differentiable arcs, except possibly at the end-points of the interval and points where the Jacobian is zero.

The points corresponding to $t=0$ are particularly important. It follows from Theorem 2 that the roots of $u_A(0,z)=0$ are double roots, and so if there are d distinct roots, we must have $h=2d$, provided that $p_0(0)\neq 0$. Consequently the images of $[0,4]$ occur in d pairs, the end-points corresponding to $t=0$ of two paired arcs being coincident. We shall refer to such a pair of arcs as a *segment*.

In terms of the parameter θ (recall that $t=4\cos^2(\theta/2)$), a segment is the image of the interval $[0,2\pi]$. It is not however a homeomorphic image of the circle $|s|=1$: since the transformation $t=s+s^{-1}+2$ is not regular at $s=0$ the image of $\theta=0$ is (in general) different from the image of $\theta=2\pi$.

In summary, the set $E(A)$ consists of d segments, together with some degenerate arcs. The end-points of each segment are two distinct roots of $u_A(4,z)=0$, and each segment contains a double root of $u_A(0,z)=0$ as an interior point. The segments are smooth, except at points where the Jacobian vanishes.

For each t the Jacobian of the mapping $z\mapsto u_A(t,z)$ is $|u'_A(t,z)|^2$, where u'_A denotes the derivative with respect to z. So the Jacobian vanishes if and only if $u'_A(t,z)=0$. Thus the condition that there is a point which lies on one of the equimodular curves, and where the Jacobian vanishes, is that there exist t^* and z^* such that $u_A(t^*,z^*)=0$ and $u'_A(t^*,z^*)=0$. In other words, there is a value t^* such that the equations $u_A(t^*,z)=0$ and $u'_A(t^*,z)=0$ have a common root.

So we have to find the values of t for which

$$u_A(t,z)=p_0(t)z^{2d}+p_1(t)z^{2d-1}+\cdots+p_{2d}(t)$$

has a double zero in z. The condition for this to happen is that the discriminant of $u_A(t,z)$ vanishes. Since the discriminant is (a multiple of) the resultant of $u_A(t,z)$ and $u'_A(t,z)$, it is a polynomial $\mathrm{disc}_A(t)$ with integer coefficients. In fact, $\mathrm{disc}_A(t)$ has a factor $(t-t_0)^\mu$ if and only if $u_A(t_0,z)=0$ has μ double roots, and in particular, since all the roots of $u_A(0,z)=0$ are double roots, this means that $\mathrm{disc}_A(t)$ has a factor t^d. Thus in order to find singularities of the equimodular curves we have to find the roots of

$$t^{-d}\,\mathrm{disc}_A(t)=0 \quad \text{for } t\in(0,4),$$

and locate the corresponding points $z\in E(A)$. Each root t_c will give rise to $2d$ points z satisfying $u_A(t_c,z)=0$, two on each of the d segments, but (in general) only one of these points is a singularity.

The general theory of singularities has been studied by Whitney [16] and others. One type of singularity that turns up in the study of zeros of chromatic polynomials is the *real-crossing singularity*. (In fact no other types of singularity have been encountered as yet.) A real-crossing singularity occurs at a point x_c on the real axis, and it can be explained as follows. An explicit example is given in Section 9.

When x is real the eigenvalues of $A(x)$ are either real numbers or pairs of conjugate complex numbers. Since conjugate complex numbers have equal modulus, it is to be expected that parts of the real axis belong to $E(A)$. The typical situation is that there is an interval $[x_1,x_2]$ and a pair of eigenvalues $\lambda(x),\bar{\lambda}(x)$, such that the ends of the interval are points where $\lambda = \bar{\lambda}$: in other words, the two eigenvalues are real and equal at x_1 and x_2. When this happens, x_1 and x_2 must be the end-points of segments of $E(A)$.

In 'regular' cases the interval $[x_1, x_2]$ is a complete segment of $E(A)$, but it is possible that there is a singularity x_c inside the interval. In this situation, the interval $[x_1, x_c]$ is part of a segment that is non-differentiable at x_c and is completed by an arc joining x_c to a complex point w. Similarly, the reverse interval $[x_2, x_c]$ is part of a segment that is non-differentiable at x_c and is completed by an arc joining x_c to $\bar{w}$. These two segments are shaped like $\rfloor$ and $\lceil$ respectively, and together they form a cross with centre at x_c.

In terms of the parameter t, the behaviour at x_c can be visualized as follows. When $t = t_c - \varepsilon$ there are roots of $u_A(t_c, z) = 0$ just above and just below x_c. When $t = t_c$, these roots collide. When $t = t_c + \varepsilon$ the roots are on the real axis, one either side of x_c.

6. Reducible matrices

In this section we discuss the equimodular curves in the case when A is *reducible*, in the sense that $A(z)$ is similar to a matrix

$$\begin{pmatrix} B(z) & V(z) \\ O & C(z) \end{pmatrix},$$

where the *constituents* $B(z)$ and $C(z)$ are square matrices and O is a matrix consisting entirely of 0's. In this situation, it is clear that the eigenvalues of $A(z)$ are those of $B(z)$ and $C(z)$. The equimodular curves are determined by two eigenvalues of $B(z)$, or two eigenvalues of $C(z)$, or one eigenvalue of $B(z)$ and one eigenvalue of $C(z)$. If we denote by $E(B, C)$ the set of points z where there is an eigenvalue of $B(z)$ and an eigenvalue of $C(z)$ with the same modulus, then

$$E(A) = E(B) \cup E(C) \cup E(B, C).$$

It will be shown that if B and C are distinct and irreducible, $E(B, C)$ is a set of closed curves. The proof is given in the algebraic framework developed in Sections 2 and 3, although it is possible that a more direct proof could be found.

If $A(z)$ is reducible, its characteristic polynomial $a(\lambda)$ is equal to $b(\lambda)c(\lambda)$, where $u(\lambda)$ and $w(\lambda)$ are the characteristic polynomials of the constituents $B(z)$ and $C(z)$. For the time being we continue to work with generic polynomials—that is, we use the indeterminate a_i instead of the function $a_i(z)$. Thus, in the reducible case, the coefficients $a_1, a_2, \dots, a_m$ of $a(\lambda)$ are given in terms of the coefficients $b_1, b_2 \dots, b_k$ of $b(\lambda)$ and the coefficients $c_1, c_2 \dots, c_\ell$ of $c(\lambda)$ by the usual rule for multiplying polynomials.

Lemma 5. *If* $a(\lambda) = b(\lambda)c(\lambda)$, *then*

$$\det R(a_s, a) = \det R(b_s, b) \det R(c_s, c) \det R(b_s, c) \det R(b, c_s).$$

Proof. This follows from the general result [7, p. 73] that for polynomials α, β, γ,

$$\det R(\alpha, \beta\gamma) = \det R(\alpha, \beta) \det R(\alpha, \gamma). \qquad \square$$

We have shown that the polynomial $\rho_m(s) = \det R(a_s, a)$ is equal to $a_m(s-1)^m \delta_m(s)$, where $\delta_m(s)$ is a reciprocal polynomial. Here it is helpful to indicate the names of the indeterminates and write $\delta_m(s)$ as $\delta_m(s, a)$. Thus

$$\det R(a_s, a) = a_m(s-1)^m \delta_m(s, a),$$

$$\det R(b_s, b) = b_k(s-1)^k \delta_k(s, b),$$

$$\det R(c_s, c) = c_\ell(s-1)^l \delta_\ell(s, c).$$

Since $m = k + \ell$ and $a_m = b_k c_\ell$ it follows from Lemma 5 that

$$\delta_m(s, a) = \delta_k(s, b)\delta_\ell(s, c) \det R(b_s, c) \det R(b, c_s).$$

Here $\det R(b_s, c)$ is a polynomial in s of degree kl, which we shall denote by

$$q(s) = \sum_{i=0}^{k\ell} \kappa_i s^{k\ell - i}.$$

Lemma 6. *With the notation as above*, $\det R(b, c_s)$ *is the reverse polynomial*

$$\tilde{q}(s) = \sum_{i=0}^{k\ell} \kappa_i s^i.$$

Proof. It is clear from the definition of the resultant that, for any constant σ,

$$\det R(b_\sigma, c_\sigma) = \sigma^{k\ell} \det R(b, c).$$

Hence

$$\det R(b, c_s) = s^{k\ell} \det R(b_{s^{-1}}, c) = s^{k\ell} \kappa(s^{-1}) = \tilde{q}(s). \qquad \square$$

The next lemma is the algebraic form of the fact that $E(A) = E(B) \cup E(C) \cup E(B, C)$.

Lemma 7. *Given* A, B, C *as in the preceding discussion*, *there is a polynomial function* $v_{B,C} : \mathbb{R} \times \mathbb{C} \to \mathbb{C}$ *such that* $v_A = v_B v_C v_{B,C}$ *and*

$$E(B, C) = \{z \in \mathbb{C} \mid v_{B,C}(t, z) = 0 \quad for\ some\ z \in [0, 4]\}.$$

Proof. Lemmas 5 and 6 imply that

$$\delta_m(s, a) = \delta_k(s, b)\delta_\ell(s, c) q(s) \tilde{q}(s).$$

Inserting the relevant functions $r_m(t), r_k(t), r_\ell(t)$ gives

$$s^{m(m-1)/2} r_m(t) = s^{k(k-1)/2} r_k(t) s^{\ell(\ell-1)/2} r_\ell(t) q(s)\tilde{q}(s).$$

Clearly $q(s)\tilde{q}(s)$ is a reciprocal polynomial of degree $2kl$ in s. Thus we can make the substitution $t = s + s^{-1} + 2$ and obtain

$$q(s)\tilde{q}(s) = s^{k\ell} q(s) q(s^{-1}) = s^{k\ell} r_{k,\ell}(t),$$

where $r_{k,\ell}(t)$ is a polynomial of degree $k\ell$ in t, and $r_m(t) = r_k(t) r_\ell(t) r_{k,l}(t)$. Replacing the indeterminates by the relevant functions, we get

$$v_A(t,z) = v_B(t,z) v_C(t,z) v_{B,C}(t,z).$$

Since we know that v_A, v_B, v_C define $E(A), E(B), E(C)$, respectively, it follows that $v_{B,C}$ defines $E(B,C)$. □

Theorem 3. *Suppose that $B(z)$ and $C(z)$ are distinct and irreducible. Then the equimodular curves $E(B,C)$ are closed curves.*

Proof. The value $t = 4$ corresponds to $s = 1$, and so $v_{B,C}(4,z)$ is obtained by substitution in the generic polynomial $r_{k,\ell}(4) = q(1)\tilde{q}(1)$. But

$$q(1)\tilde{q}(1) = (\det R(b,c))^2.$$

Replacing the indeterminates by the relevant functions, it follows that there is a polynomial $g_{B,C}$ such that

$$v_{B,C}(4,z) = g_{B,C}(z)^2.$$

Hence the roots of $v_{B,C}(4,z) = 0$ are double roots, and the segments comprising $E(B,C)$ link up to form closed curves. □

7. The dominance property

The intended application of the work presented here concerns the limit set of the zeros of certain sequences of polynomials. A theorem of Beraha–Kahane–Weiss [1] asserts that (apart possibly from some isolated points) the limit points are a subset of those parts of the equimodular curves that have a 'dominance' property, which we now define.

For each $z \in \mathbb{C}$ the *spectral radius* of the square matrix $A(z)$ is

$$m_A(z) = \max\{|\lambda| \mid \det(\lambda I - A(z)) = 0\}.$$

We say that a point ζ is *dominant* for A if there are two eigenvalues λ_1, λ_2 of $A(\zeta)$ such that

$$|\lambda_1| = |\lambda_2| = m_A(\zeta).$$

By convention, this includes the case where there is an eigenvalue λ, of algebraic multiplicity 2 or more, such that $|\lambda| = m_A(\zeta)$. We shall denote the set of dominant points for A by $D(A)$. Clearly $D(A) \subseteq E(A)$.

Points that lie on an equimodular curve are not necessarily dominant, so $D(A)$ is, in general, a proper subset of $E(A)$. Roughly speaking, if $A(z)$ is an $m \times m$ matrix there are $\frac{1}{2}m(m-1)$ equimodular curves, only one of which is dominant. Thus any method of determining $D(A)$ which involves finding $E(A)$ before applying the dominance condition, is not very efficient.

A better approach is based on the observation (Section 5) that every equimodular curve is the union of segments. Each segment has end-points given by a root of $v_A(4,z) = 0$, and contains a double root of $v_A(0,z) = 0$. We shall refer to these points as *special points*. For each special point σ it is easy to determine whether or not σ has the dominance property, by explicitly computing all the eigenvalues of $A(\sigma)$. Then, for each dominant special point we may carry out a local search to determine the behaviour of the of equimodular curve in its vicinity. The results of Salas and Sokal [13, Section 4.2] are useful here. This process can be repeated, and it will (in favourable cases) produce a dominant equimodular curve. But there can be complications caused by singularities (Section 5) and triple points (see below, Section 8).

Now consider the case where $A(z)$ is reducible, with just two distinct irreducible constituents, $B(z)$ and $C(z)$. We shall say that a point z^* is *dominant for the pair* (B, C) if $m_B(z^*) = m_C(z^*)$, and denote the set of points dominant for (B, C) by $D(B, C)$. Clearly, $D(B, C)$ is a subset of $E(B, C)$ and

$$D(B, C) \subseteq D(A) \subseteq D(A) \cup D(B) \cup D(B, C).$$

According to Lemma 7, the curves comprising $E(B, C)$ are defined by the vanishing of a polynomial function $v_{B,C}(t, z)$. As in the irreducible case these curves can be decomposed into segments, and here we have the additional property that the special points defined by $v_{B,C}(4, z) = 0$ occur in pairs, so that the segments link up to form closed curves. Since $v_{B,C}(t, z)$ is of the form $q(s, z)\tilde{q}(s, z)$, and the values $t = 4$ and 0 correspond to $s = 1$ and -1 respectively, it is easier to work with the polynomial q. Thus the local search method outlined above can be used to find $D(B, C)$, but the same difficulties may arise.

8. Triple points

An equimodular curve is smooth, except at points where a Jacobian vanishes However, there may be points where the curve is smooth but the dominance property is not preserved. Indeed, the dominance property will be altered at points where a third eigenvalue is equal in modulus to the two eigenvalues that define the curve. We shall say that τ is a *triple point* if three (or more) eigenvalues have equal modulus at τ. (Salas and Sokal [13] refer to this as a T-point.) A triple point lies on three equimodular curves, corresponding to the three possible pairs of these three eigenvalues.

The method outlined in the previous section will recognize a triple point. As we approach a triple point, the local search will encounter extra points having the equimodular and/or dominance properties. This also happens as we approach a singularity, but a triple point can be distinguished from a singularity by testing whether the Jacobian is zero.

It is possible that all three eigenvalues involved in a triple point are dominated by some other eigenvalue, in which case the triple point plays no part in the determination of $D(A)$. But if a triple point τ lies on a curve that is known to be dominant, each of the three curves passing through τ has the property that its points on one side of τ are in $D(A)$, while those on the other side are not.

When A is reducible, with two irreducible constituents B and C, it is natural to begin by finding $D(B)$ and $D(C)$. Only a part of $D(B)\cup D(C)$ is in $D(A)$: a point ζ in $D(B)$ is dominant for A if and only if the two dominant equimodular eigenvalues of B dominate *all* the eigenvalues of C at ζ; and similarly with B and C switched.

The application of this simple criterion may involve the determination of triple points. The following useful theorem was noted by Matveev and Shrock [11]. It implies that, in the construction of $D(B,C)$ by the local search process, triple points will occur if and only if the curve under construction hits $D(B)$ or $D(C)$, and such a triple point will separate a part of $D(B)$ or $D(C)$ that belongs to $D(A)$ from a part that does not.

Theorem 4. *A triple point that belongs to $D(B,C)$ must belong to $D(B)\cup D(C)$.*

Proof. Consider part of an equimodular curve that belongs to $D(B,C)$. With suitable care about the domain of definition, we may suppose that there are eigenvalues $\lambda_1(z), \mu_1(z)$ such that the curve is defined by an equation of the form $|\lambda_1(z)| = |\mu_1(z)|$, where $|\lambda_1(z)| = m_B(z)$ and $|\mu_1(z)| = m_C(z)$. Then at a triple point τ there is a third eigenvalue equal in modulus to $\lambda_1(\tau)$ and $\mu_1(\tau)$. Without loss of generality, we may take it to be an eigenvalue $\mu_2(\tau)$ of $B(\tau)$. It follows that the two other equimodular curves passing through τ are defined by the equations

$$|\lambda_1(z)| = |\mu_2(z)|, \quad |\mu_1(z)| = |\mu_2(z)|.$$

Since $|\mu_1(z)| = m_C(z)$ in a neighbourhood of τ, it follows that the second curve is in $D(C)$, and hence τ is in $D(C)$. $\square$

9. An example

Let

$$B(z) = \begin{pmatrix} z^4+2z^3+3z^2+z+1 & z^3+z & z^3+z^2+2z \\ -1 & z^2 & -z \\ -(z^2+z+1) & -z & -z \end{pmatrix},$$

$$C(z)=\begin{pmatrix} -z & 1 & 0 & -1 \\ -z & z^2+1 & 1 & z \\ 1 & -z+1 & -z & -1 \\ -1 & z & 0 & 1 \end{pmatrix}.$$

These matrices arise in the discussion of the limit points of the chromatic roots of the family of generalized dodecahedra [4,5]. They do not completely solve that problem (another matrix is also involved); however, their individual properties and their interaction as constituents of a matrix $A(z)$ provide a good illustration of the methods developed above.

The coefficients of the characteristic polynomials of $B(z)$ and $C(z)$ are:

$$b_1(z) = -z^4 - 2z^3 - 4z^2 - 1,$$

$$b_2(z) = z(z+1)(z^4 + z^3 + 2z^2 + 2),$$

$$b_3(z) = -z^4 - 2z^3 - z^2,$$

$$c_1(z) = -z^2 + 2z - 2,$$

$$c_2(z) = -2z^3 + z^2 - 2z - 1,$$

$$c_3(z) = -z^4 + 1,$$

$$c_4(z) = z^2 + +2z + 1.$$

The polynomials v_B and v_C can be calculated by substituting these coefficients in the generic polynomials $r_3(t)$ and $r_4(t)$. For v_B, the result is of the form

$$v_B(t,z) = -z^2(z+1)^2(z^{16} + 6z^{15} + (25-t)z^{14} + \cdots),$$

and from it the discriminant $\mathrm{disc}_B(t)$ can be obtained.

The form of $v_B(t,z)$ indicates that there are two degenerate arcs, 0 and -1, and the rest of $E(B)$ comprises 8 segments. The special points are listed in [4]. The dominance condition eliminates all of them except:

$$(t=4): -1.8726 \pm 1.1275i, \quad -0.3412 \pm 1.1615i, \quad 0.1541, \quad 0.6066;$$
$$(t=0): -1.0788 \pm 1.7292i, \quad 0.1601 \pm 0.4718i.$$

Here, and in the ensuing discussions, points are represented by an approximation to four decimal places.

All the 4-points except the second pair are single roots. Thus $-1.8726 + 1.1275i$ is the end point of a segment of a dominant equimodular curve. This segment passes through the 0-point $1.0788 + 1.7292i$ and terminates at the 4-point $-0.3412 + 1.1615i$. Since this is a double root, another segment begins at this point; it passes through the 0-point $0.1601 + 0.4718i$ and continues until it hits the real axis at 0.3369. This point is a real-crossing singularity, of the type described in Section 5. In fact there are only two roots of $\mathrm{disc}_B(t)=0$ in the open interval $(0,4)$, 2.3587 and 3.0850, and the first one determines the singularity at 0.3369. (The second one also corresponds to a real-crossing singularity, but it is on a non-dominant curve [4].) Thus the segment approaching 0.3369 from above is non-differentiable at this point: we may think of it as turning sharp right along the real axis, and terminating at 0.1541.

Starting from $-1.8726-1.1275i$ we obtain the conjugates of the points just described, except that the segment approaching 0.3369 from below turns sharp right at that point and terminates at 0.6066. All the dominant special points have been covered, so we conclude that $D(B)$ is the union of four segments, two of which are non-differentiable at the singularity.

For the matrix $C(z)$, we have

$$v_C(t,z) = -(z+1)^2((t-4)z^{16} + (2t^2-12t+8)z^{15} + (t^3-8t^2+28t-18)z^{14} + \cdots).$$

Here too only a few of the special points have the dominance property:

$$(t=4):\ -1.5684 \pm 2.1597i, \quad 0.5000;$$

$$(t=0):\ 0.5324 \pm 1.5856i.$$

The point 0.5000 is a double root.

Let Γ_1 be the segment with end-point $-1.5684+2.1597i$. This segment passes through the 0-point $0.5324+1.5856i$ and remains dominant until it hits a triple point α at $0.5872+1.4516i$. The two other segments Γ_2 and Γ_3 passing through α intersect Γ_1 again at another triple point $\beta = 0.5944+1.2671i$. Between these two points Γ_2 and Γ_3 are dominant, but Γ_1 is not. However Γ_1 becomes dominant again after passing through β, and it terminates at 0.5000. The rest of $D(C)$ is conjugate to what has just been described: thus it comprises the entire segment $\bar{\Gamma}_1$, except for a small section between $\bar{\alpha}$ and $\bar{\beta}$, where parts of the segments $\bar{\Gamma}_2$ and $\bar{\Gamma}_3$ are dominant.

It remains to find $D(B,C)$. The equimodular curves $E(B,C)$ are obtained from the generic polynomial $r_{3,4}(t) = s^{-12}q(s)\tilde{q}(s)$, where

$$q(s) = \det \begin{pmatrix} s^3 & s^2b_1 & sb_2 & b_3 & 0 & 0 & 0 \\ 0 & s^3 & s^2b_1 & sb_2 & b_3 & 0 & 0 \\ 0 & 0 & s^3 & s^2b_1 & sb_2 & b_3 & 0 \\ 0 & 0 & 0 & s^3 & s^2b_1 & sb_2 & b_3 \\ 1 & c_1 & c_2 & c_3 & c_4 & 0 & 0 \\ 0 & 1 & c_1 & c_2 & c_3 & c_4 & 0 \\ 0 & 0 & 1 & c_1 & c_2 & c_3 & c_4 \end{pmatrix}.$$

Thus q is a polynomial of degree 12 in s, $\sum \kappa_i s^{12-i}$. The coefficients κ_i are integral linear combinations of terms of the form $\psi\omega$, where each ψ is a monomial of weight

i in the b's and each ω is a monomial of weight $12-i$ in the c's. For example,

$$\kappa_0 = c_4^3, \quad \kappa_1 = -b_1 c_3 c_4^2, \quad \kappa_2 = b_2 c_3^2 c_4 + b_1^2 c_2 c_4^2 - 2 b_2 c_2 c_4^2.$$

Substituting the relevant functions of z, as given above, we obtain a polynomial function of s and z,

$$q(s,z) = (z+1)^4 q_0(s,z),$$

where q_0 is a polynomial of degree 22 in z. Putting $s=1$ the coefficients of z^{22}, z^{21} and z^{20} vanish, and we get

$$q_0(1,z) = z(z+1)^4(z^5 + 3z^3 + 2z - 2)\, p_9(z),$$

where

$$p_9(z) = 4z^9 + 6z^8 + 10z^7 + 9z^6 - 12z^5 + 6z^4 - 28z^3 + 15z^2 - 6z + 4.$$

Similarly putting $s=-1$ we get

$$\begin{aligned} q_0(-1,z) = {} & 4z^{22} + 12z^{21} + 42z^{20} + 48z^{19} + 126z^{18} + 42z^{17} + 233z^{16} \\ & - 226z^{15} + 351z^{14} - 642z^{13} + 852z^{12} - 1038z^{11} + 1476z^{10} \\ & + 1010z^9 + 1107z^8 - 1010z^7 + 859z^6 + -670z^5 + 380z^4 \\ & - 206z^3 + 88z^2 - 20z + 2. \end{aligned}$$

It turns out that only 4 of the roots of $q(1,z)=0$ and 2 of the roots of $q(-1,z)=0$ have the dominance property. (It follows from the general theory that these points are double roots of $v_{B,C}(4,z)=0$ and $v_{B,C}(0,z)=0$ respectively.) They are:

$$(s=1) : -1.0000, \quad -0.4660 \pm 1.4456i, \quad 0.6383;$$

$$(s=-1) : 0.2574 \pm 0.6675i.$$

In particular, the point 0.6383 is the largest real root of $p_9(z)=0$. This identifies the critical value 2.6383 found by Chang [5], a value which has special significance in the physical context.

In the construction of $D(B,C)$ four triple points are encountered. Starting from 0.6383, two segments extend to the left until they hit triple points on $D(C)$ at $0.5043 \pm 0.1927i$. Here dominance is acquired by segments that pass through the dominant special points at $0.2574 \pm 0.6675i$ and terminate at $-0.4660 \pm 1.4456i$. The end points are double roots, so new segments start there, hitting $D(B)$ at triple points $-0.6735 \pm 1.5822i$. At these points dominance is acquired by another pair of segments that join up at the dominant special point -1.0000.

Let $A(z)$ be a matrix with constituents $B(z)$ and $C(z)$. The preceding description shows that the subset $D(B,C)$ of $D(A)$ is the union of parts of three closed curves in $E(B,C)$. It also determines the parts of $D(B)$ and $D(C)$ that belong to $D(A)$: they are the parts of $D(B)$ to the left of the triple points that lie on it, and the part of $D(C)$ that joins the triple points that lie on it.

References

[1] S. Beraha, J. Kahane, N.J. Weiss, Limits of zeros of recursively defined families of polynomials, Studies in foundations and combinatorics, Adv. Math. Suppl. Studies 1 (1978) 213–232.

[2] N.L. Biggs, A matrix method for chromatic polynomials, J. Combin. Theory (B) 82 (2001) 19–29.

[3] N.L. Biggs, R.M. Damerell, A.D. Sands, Recursive families of graphs, J. Combin. Theory (B) 12 (1972) 123–131.

[4] N.L. Biggs, P. Reinfeld, The chromatic roots of generalised dodecahedra, CDAM Research Report Series, LSE-CDAM-2000-7, June 2000.

[5] S.-C. Chang, Chromatic polynomials for lattice strips with cyclic boundary conditions, Physica A 296 (2001) 495–522.

[6] S.-C. Chang, R. Shrock, Ground state entropy of the Potts antiferromagnet on strips of the square lattice. http://uk.arXiv.org/ps/cond-mat/0004161

[7] D. Cox, J. Little, D. O'Shea, Using Algebraic Geometry, Springer, New York, 1998.

[8] T.D. Lee, C.N. Yang, Statistical theory of equations of state and phase transitions II. Lattice gas and Ising model, Phys. Rev. 87 (1952) 410–419.

[9] E.H. Lieb, A.D. Sokal, A general Lee-Yang theorem for one-component and multicomponent ferromagnets, Commun. Math. Phys. 80 (1981) 153–179.

[10] P. Martin, Analytic properties of the partition function for statistical mechanical models, J. Phys. A 19 (1986) 3267–3277.

[11] V. Matveev, R. Shrock, Complex-temperature phase diagram of the 1D Z_6 clock model and its connection with higher-dimensional models, Phys. Lett. A 221 (1996) 343–349.

[12] R. Read, G.F. Royle, Chromatic roots of families of graphs, in: Y. Alavi, G. Chartrand, O.R. Ollerman, A.J. Schwenk (Eds.), Graph Theory, Combinatorics and Applications, Wiley, New York, 1991.

[13] J. Salas, A.D. Sokal, Transfer matrices and partition-function zeros for antiferromagnetic Potts models. http://uk.arXiv.org/ps/cond-mat/0004330

[14] R. Shrock, S.-H. Tsai, Asymptotic limits and zeros of chromatic polynomials and ground state entropy of Potts antiferromagnets, Phys. Rev. E55 (1997) 5165–5179.

[15] R. Shrock, S.-H. Tsai, Ground-state degeneracy of Potts antiferromagnets: homeomorphism classes with non-compact W boundaries, Physica A 265 (1999) 186–223.

[16] H. Whitney, Elementary structure of real analytic varieties, Ann. Math. 66 (1957) 545–556.

[17] D. Wood, The exact location of partition function zeros, a new method for statistical mechanics, J. Phys. A 18 (1985) L917–L921.