Equimodular curves for reducible matrices

Norman Biggs

Centre for Discrete and Applicable Mathematics
London School of Economics
Houghton Street
London WC2A 2AE
U.K.

CDAM Research Report LSE-CDAM-2001-01

January 2001

Abstract

This paper is a continuation of *Equimodular curves*, LSE-CDAM-2000-17. First, it is shown by algebraic means that the equimodular curves for a reducible matrix are closed curves. Then the question of dominance is investigated, and a method of constructing the dominant equimodular curves for a reducible matrix is suggested. The method is illustrated by examples that arise in the calculation of the chromatic polynomials of an interesting family of graphs.

MSC 2000: 05C15, 05C50, 05E99.

Equimodular curves for reducible matrices

1. Introduction

This paper is a continuation of [BC], and the general motivation for it will be found in the introduction to that paper. The notation and results of [BC] are assumed, although some modifications to the notation will be made.

The objects under discussion are functions F from $\mathbf{C}$ to a ring of square matrices over $\mathbf{C}$, with the property that each component of F is a polynomial function $\mathbf{C} \to \mathbf{C}$. We shall denote by $E(F)$ the *equimodular curves* for F, that is, the set of points z for which $F(z)$ has two eigenvalues with the same modulus. Conventionally, this includes the points where $F(z)$ has an eigenvalue of algebraic multiplicity 2 or more.

If F, G are two such functions (possibly with different sizes) we denote by $E(F, G)$ the set of points z where there is an eigenvalue of $F(z)$ and an eigenvalue of $G(z)$ with the same modulus.

The main results obtained here concern the equimodular curves in the case when F is *reducible*, in the sense that $F(z)$ is similar to a matrix

$$\begin{pmatrix} U(z) & V(z) \\ O & W(z) \end{pmatrix},$$

where the *constituents* $U(z)$ and $W(z)$ are square matrices and O is a matrix consisting entirely of 0's. In this situation, it is clear that the eigenvalues of $F(z)$ are those of $U(z)$ and $W(z)$. The equimodular curves are determined by two eigenvalues of $U(z)$, or two eigenvalues of $W(z)$, or one eigenvalue of $U(z)$ and one eigenvalue of $W(z)$. That is,

$$E(F) = E(U) \cup E(W) \cup E(U, W).$$

In the first part of the paper it will be shown that, if U and W are distinct and irreducible, $E(U, W)$ is a set of closed curves. The proof is given in the algebraic framework of [BC], although it is possible that a more direct proof could be found.

For the intended application of these results, we are particularly interested in the subset $D(F)$ of $E(F)$ containing points that are 'dominant' in a certain sense (the definition is given in Section 5). Roughly speaking, when $F(z)$ is an $m \times m$ matrix we may expect $\frac{1}{2}m(m-1)$ equimodular curves, but only one of them will be in $D(F)$. In the second part of the paper we investigate the structure of the set $D(F)$, and suggest a method of computing it. We show that, when F is reducible, the dominant curves can be analyzed by a method that makes use of this property. The method is illustrated by examples involving matrices that arise in the calculation of the chromatic roots of an interesting family of graphs.

2. A family of examples

The family of examples discussed in this section is very simple. But it does illustrate the various situations that can arise, including the reducible case. More substantial examples will be discussed in Sections 5, 6, and 7.

For each integer j let

$$X_j(z) = \begin{pmatrix} 4z & 2z+j \\ 2 & 2z \end{pmatrix}.$$

The eigenvalues of $X_j(z)$ are the roots of $x_j(\lambda) = 0$, where

$$x_j(\lambda) = \lambda^2 - 6z\lambda + (8z^2 - 4z - 2j).$$

According to [BC] there is a polynomial $v_j(t,z)$ such that the points $z \in \mathbf{C}$ where the eigenvalues have equal modulus, are those for which $v_j(t,z) = 0$ for some t in the range $0 \le t \le 4$. For a quadratic polynomial $v_j(t,z)$ is obtained by substituting the explicit forms of the coefficients, in this case

$$a_1(z) = -6z, \quad a_2(z) = 8z^2 - 4z - 2j$$

in the expression $ta_2 - a_1^2$. Hence

$$v_j(t,z) = (8t - 36)z^2 - 4tz - 2jt.$$

The set $E(X_j)$ can now be determined, using the method described in [BC, Section 6]. Here there is just one continuous arc, containing the unique (double) root of the equation $v_j(0,z) = 0$, and this is the point 0 for all j. The endpoints of the arc are the roots of $v_j(4,z) = 0$, that is, the points

$$\sigma_j = -2 - \sqrt{4 - 2j}, \quad \tau_j = -2 + \sqrt{4 - 2j}.$$

The arc may not be differentiable at internal points where the discriminant $\mathrm{disc}_j(t)$ of v_j (regarded as a polynomial in z) vanishes; that is

$$\mathrm{disc}_j(t) \;=\; (64j + 16)t^2 - 288jt = 0.$$

This happens when $t = t_j = 18j/(4j+1)$, a value which lies in the range $0 < t < 4$ only when $j = 1$.

The sets $E(X_j)$ in the various cases are discussed in detail below and illustrated in Figure 1.

Case 1: $j < 0$ Here the endpoints σ_j, τ_j are real, and $\sigma_j < -4$, $\tau_j > 0$. For all points $z = \zeta$ on the real axis such that $\sigma_j < \zeta < \tau_j$, the roots of $x_j(\lambda) = 0$ are complex conjugates, and hence equal in modulus. It follows that the equimodular curve is the interval $[\sigma_j, \tau_j]$ of the real axis.

Case 2: $j = 0$ The only difference from the previous case is that $\tau_0 = 0$, and the 'right half' of the equimodular curve collapses to the single point 0. The entire curve is the interval $[-4, 0]$.

Figure 1: the sets $E(X_j)$ in the various cases.

Case 1: $j < 0$

Case 2: $j = 0$

Case 3: $j = 1$

Case 4: $j = 2$

Case 5: $j > 2$

Case 3: $j = 1$ Here the endpoints are $\sigma_1 = -2 - \sqrt{2}$ and $\tau_1 = -2 + \sqrt{2}$. For real values $z = \zeta$ such that $\sigma_1 < \zeta < \tau_1$ the eigenvalues are complex conjugates, and so the interval $[\sigma_1, \tau_1]$ is part of the equimodular curve. However, in this case it cannot be the entire curve, because the point 0 is certainly on the curve, but not in the interval. The mystery is solved by the observation that when $j = 1$ the discriminant $\text{disc}_1(t)$ vanishes at the value $t_1 = 18/5$. For t in the range $0 < t < t_1$ we have $\text{disc}_1(t) < 0$, so the equation $v_1(t, z)$ has complex conjugate roots. The equation $v_1(t_1, z) = 0$ has a double root -1, and the arc has a self-intersection at that point.

We can visualise the behaviour in the neighbourhood of -1 as follows. For $t = t_1 - \epsilon$ the two roots of $v_1(t, z) = 0$ are complex conjugates lying just above and just below -1. At $t = t_1$ the roots collide. For $t = t_1 + \epsilon$ the roots lie on the real axis, on either side of -1.

Case 4: $j = 2$ In this case the characteristic polynomial $x_2(\lambda)$ factorises, and the eigenvalues are $4z + 2$ and $2z - 2$. This means that $X_2(z)$ is reducible, with constituents the 1×1 matrices $[4z + 2]$ and $[2z - 2]$.

The endpoints σ_2 and τ_2 coincide at -2, so the arc is a closed curve, the circle defined by $|4z + 2| = |2z - 2|$.

Case 5: $j > 2$ In this case σ_j and τ_j are complex conjugates, so the arc is an open arc.

3. The polynomial criterion for reducible matrices

In this section we extend the general theory developed in Section 2 of [BC]. Given an $m \times m$ matrix $F(z)$, each of whose entries is a polynomial function of the complex variable z, with integer coefficients, denote by $f(\lambda)$ be the characteristic polynomial of $F(z)$, and let $f_i(z)$ be the coefficient of λ^{m-i} in $f(\lambda)$, that is,

$$f(\lambda) = \det(\lambda I - F(z)) = \lambda^m + f_1(z)\lambda^{m-1} + f_2(z)\lambda^{m-2} + \cdots + f_m(z).$$

The coefficients $f_i(z)$ are the sums of principal minors of $F(z)$, and so they are polynomials with integer coefficients. Let

$$f_s(\lambda) = f(s\lambda) = s^m\lambda^m + s^{m-1}f_1(z)\lambda^{m-1} + s^{m-2}f_2(z)\lambda^{m-2} + \cdots + f_m(z).$$

The condition that f has two roots with equal modulus is equivalent to the condition that f and f_s have a common root for some s such that $|s| = 1$. This happens when the *resultant* $\det R$ vanishes, where $R = R(f_s, f)$ is a $2m \times 2m$ matrix whose entries depend on the coefficients of f and f_s, that is, s and $f_1(z), f_2(z), \ldots, f_m(z)$. Thus, $F(z)$ has two eigenvalues with equal moduli if and only if the resultant has a zero on the unit circle $|s| = 1$.

If $F(z)$ is reducible, its characteristic polynomial $f(\lambda)$ is equal to $u(\lambda)w(\lambda)$, where $u(\lambda)$ and $w(\lambda)$ are the characteristic polynomials of the constituents $U(z)$ and $W(z)$. For the time being it is convenient to work with generic polynomials - that is, we use a symbol f_i instead of the function $f_i(z)$. In the reducible case, the coefficients $f_1, f_2, \ldots, f_m$ of f are given in terms of the coefficients $u_1, u_2 \ldots, u_k$ of u and the coefficients $w_1, w_2 \ldots, w_\ell$ of w by the usual rule for multiplying polynomials.

Lemma 1 If $f(\lambda) = u(\lambda)w(\lambda)$, then

$$\det R(f_s, f) = \det R(u_s, u) \det R(w_s, w) \det R(u_s, w) \det R(u, w_s).$$

Proof This follows from the general result [6, p.73] that for polynomials α, β, γ,

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

$\square$

It is shown in [BC] that the polynomial $\det R(f_s, f)$ is equal to $f_m(s-1)^m \Delta_m(s, f)$. Here $\Delta_m(s, f)$ is a reciprocal polynomial of degree $m(m-1)$ in s, the coefficients being linear combinations of monomials in $f_1, f_2, \ldots, f_m$. Applying this result also to u and w we have

$$\det R(f_s, f) = f_m(s-1)^m \Delta_m(s, f)$$
$$\det R(u_s, u) = u_k(s-1)^k \Delta_k(s, u)$$
$$\det R(w_s, w) = w_\ell(s-1)^l \Delta_\ell(s, w).$$

Since $m = k + \ell$ and $f_m = u_k w_\ell$ it follows from Lemma 1 that

$$\Delta_m(s, f) = \Delta_k(s, u)\Delta_\ell(s, w) \det R(u_s, w) \det R(u, w_s).$$

Lemma 2 Put $n = kl$, and let $\det R(u_s, w)$ be the polynomial

$$q(s) = \sum_{i=0}^{n} q_i s^{n-i}.$$

Then $\det R(u, w_s)$ is the reverse polynomial

$$\tilde{q}(s) = \sum_{i=0}^{n} q_{n-i} s^{n-i}.$$

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

$$\det R(u_\sigma, w_\sigma) = \sigma^n \det R(u, w).$$

Hence

$$\det R(u, w_s) = s^n \det R(u_{s^{-1}}, w) = s^n q(s^{-1}) = \tilde{q}(s).$$

$\square$

Clearly $q(s)\tilde{q}(s)$ is a reciprocal polynomial of degree $2n = 2kl$ in s, and its coefficients are linear combinations of monomials in the coefficients of u and w. Thus we can make the substitution $t = s + s^{-1} + 2$ and obtain

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

where $r_{k,\ell}(t)$ is a polynomial of degree $k\ell$ in t.

Example Suppose that $f(\lambda)$ is a cubic polynomial that factors as $u(\lambda)w(\lambda)$, where

$$u(\lambda) = \lambda^2 + u_1\lambda + u_2, \quad w(\lambda) = \lambda + w_1.$$

This means that the coefficients of $f(\lambda) = \lambda^3 + f_1\lambda^2 + f_2\lambda + f_3$ are given by

$$f_1 = u_1 + w_1, \quad f_2 = u_2 + u_1w_1, \quad f_3 = u_2w_1.$$

Using the definition of the resultant, we have

$$\det R(u_s, u) = u_2(s-1)^2\Big(u_2(s+1)^2 - u_1^2 s\Big), \quad \det R(w_s, w) = w_1(s-1),$$

$$\det R(u_s, w) = q(s) = w_1^2s^2 - u_1w_1s + w_2,$$

$$\det R(u, w_s) = \tilde{q}(s) = u_2s^2 - u_1w_1s + w_1^2.$$

The lemma tells us that $\det R(f_s, f)$ is the product of these four resultants. This is a polynomial of the form

$$u_2w_1(s-1)^3h(s) = f_3(s-1)^3h(s),$$

where

$$h(s) = (u_2(s+1)^2 - u_1^2s)\,(w_1^2s^2 - u_1w_1s + u_2)\,(u_2s^2 - u_1w_1s + w_1^2).$$

Using the equations for the coefficients f_1, f_2, f_3, it can be verified that $h(s)$ is the generic polynomial $\Delta_3(s, f)$, as given in [BC]:

$$\begin{aligned}\Delta_3(s,f) = {} & f_3^2(s^6+1) + (3f_3^2 - f_1f_2f_3)(s^5+s) \\ & + (6f_3^2 - 5f_1f_2f_3 + f_2^3 + f_1^3f_3)(s^4+s^2) \\ & + (7f_3^2 - 6f_1f_2f_3 + 2f_2^3 + 2f_1^3f_3 - f_1^2f_2^2)s^3.\end{aligned}$$

As in [BC], putting $t = s + s^{-1} + 2$ we can write $\Delta_3(s, f)$ as $s^3r_3(t)$, for a generic polynomial $r_3(t)$.

In the same way,

$$q(s)\tilde{q}(s) = u_2w_1^2(s^4+1) - (u_1w_1^3 + u_1u_2w_1)(s^3+s) + (u_1^2w_1^2 + u_2^2 + w_1^4),$$

and this can be written as $s^2r_{2,1}(t)$, where $t = s + s^{-1} + 2$ and

$$r_{2,1}(t) = u_2w_1^2t^2 - (u_1w_1^3 + u_1u_2w_1 + 4u_2w_1^2)t + (u_2 + u_1w_1 + w_1^2)^2.$$

□

4. The equimodular curves as unions of arcs

The results obtained in the preceding example are particular cases of the following. With the substitution $t = s + s^{-1} + 2$,

$$\Delta_m(s, f) = s^{m(m-1)/2}\, r_m(t),$$

where r_m is a polynomial of degree $m(m-1)/2$ in t. And similarly,

$$q(s)\tilde{q}(s) = s^{kl}\, r_{k,l}(t),$$

where $r_{k,l}$ is a polynomial of degree kl in t.

At this point, we return to the situation when f is the characteristic polynomial of a matrix $F(z)$, so that its coefficients are complex functions $f_i(z)$. Define $v_f : \mathbf{R}\times\mathbf{C} \to \mathbf{C}$ by the rule that $v_f(t, z)$ is obtained from $r_m(t)$ by replacing f_i by $f_i(z)$, and define $v_{u,w} : \mathbf{R} \times \mathbf{C} \to \mathbf{C}$ similarly, using $r_{k,l}(t)$.

Recall that when s is replaced by $t = s + s^{-1} + 2$, the condition $|s| = 1$ implies that t is real and lies in the range $0 \leq t \leq 4$. This observation leads to the main result of [BC], that the set of points z where $F(z)$ has two eigenvalues with the same modulus is given by

$$E(F) \;=\; \{z \in \mathbf{C} \mid v_f(t, z) = 0 \text{ for some } t,\ 0 \leq t \leq 4\}.$$

Except (possibly) for some isolated points, $E(F)$ is a union of homeomorphic images of the interval $[0, 4]$. All roots of $v_f(0, z) = 0$ are double roots, and consequently the images of $[0, 4]$ occur in pairs, the end-points corresponding to $t = 0$ of two paired arcs being coincident. It is convenient to use the word *segment* to denote two images of $[0, 4]$ that are paired in this way.

Thus $E(F)$ consists of a number of segments, such that the end-points of each segment are roots of $v_f(4, z) = 0$, and each segment contains a double root of $v_f(0, z) = 0$. The segments are smooth, except at points where the Jacobian vanishes.

For our present purposes, the crucial observation is that some of the roots of $v_f(4, z) = 0$ may also be double roots, in which case the segments join up. This happens, for example, in Case 4, Section 2, where the matrix $X_2(z)$ is reducible: there is just one segment, and its endpoints coincide, so that it defines a closed curve.

In general, if F is reducible with constituents U and W we have $E(F) = E(U) \cup E(W) \cup E(U, W)$. The algebraic form of this result can be deduced from the theory developed above, and it leads to the following result.

Theorem 1 Suppose that $F(z)$ is a reducible matrix with constituents $U(z)$ and $W(z)$, which are distinct and irreducible. Then the equimodular curves in $E(U, W)$ are closed curves.

Proof If $F(z)$ is reducible, its characteristic polynomial f is the product uw of the characteristic polynomials of $U(z)$ and $W(z)$. The following result is a consequence of Lemmas 1 and 2:

$$\Delta_m(s, f) = \Delta_k(s, u)\Delta_\ell(s, w)q(s)\tilde{q}(s).$$

Making the substitution $t = s + s^{-1} + 2$ and replacing the coefficients of the generic polynomials by the appropriate functions, we obtain

$$v_f(t,z) \;=\; v_u(t,z)v_w(t,z)\; v_{u,w}(t,z).$$

This is the algebraic form of the decomposition $E(F) = E(U) \cup E(W) \cup E(U,W)$. In other words

$$E(U,W) \;=\; \{z \in \mathbf{C} \mid v_{u,w}(t,z) = 0 \text{ for some } t,\; 0 \leq t \leq 4\}.$$

In order to show that $E(U,W)$ consists of closed curves, note that the value $t = 4$ corresponds to $s = 1$, and so $v_{u,w}(4,z)$ is obtained by substitution in the generic polynomial $q(1)\tilde{q}(1)$. But

$$q(1)\tilde{q}(1) \;=\; (\det R(u,w))^2.$$

It follows that all roots of $v_{u,w}(4,z) = 0$ are double roots. Hence the segments comprising $E(U,W)$ link up to form closed curves. □

5. 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 $z \in \mathbf{C}$ the *spectral radius* of the square matrix $F(z)$ is

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

We say that a point z^* is *dominant* for F if there are two eigenvalues λ_a, λ_b of $F(z^*)$ such that

$$|\lambda_a| = |\lambda_b| = m_F(z^*).$$

By convention, this includes the case where there is an eigenvalue λ_a, of algebraic multiplicity 2 or more, such that $|\lambda_a| = m_F(z^*)$. We shall denote the set of dominant points for F by $D(F)$.

Points that lie on an equimodular curve are not necessarily dominant, so $D(F)$ is, in general, a proper subset of $E(F)$. Roughly speaking, if $F(z)$ is an $m \times m$ matrix there are $\frac{1}{2}m(m-1)$ equimodular curves, only one of which is dominant. Thus the method of determining $D(F)$ used in [4], which involves finding $E(F)$ before applying the dominance condition, is not very efficient. The aim is to design a more efficient method, using the dominance condition from the outset.

A first attempt to do this is based on the idea of locating a small number of points that lie on the equimodular curves. In the terminology introduced in Section 4, every equimodular curve is the union of *segments*. Each segment has end-points given by the value $t = 4$ (equivalently $s = 1$), and contains a point given by the value $t = 0$ ($s = -1$). We shall refer to these points as *special points*. Algebraically, a special point is a root of one of the equations

$$v(4,z)=0, \quad v(0,z)=0.$$

(From now on we drop the subscripts on the v polynomials.)

For each special point z' it is easy to determine whether or not z' has the dominance property, by explicitly computing all the eigenvalues of $F(z')$. However, it is worth noting that a special point may be dominant, even though it arises from the equimodularity of two non-dominant eigenvalues. In other words, it is possible that a root of $v(4,z)=0$, which lies on a dominant equimodular curve, is not the end-point of a segment of that curve.

There now follows the outline of a suggested method for determining $D(F)$.

(i) Substitute the functional forms of the coefficients in the generic polynomial $r_m(t)$, obtaining a polynomial $v(t,z)$.
(ii) Compute the special points - that is, the roots of $v(0,z)=0$ and $v(4,z)=0$.
(iii) For each special point z', calculate the eigenvalues of $F(z')$ and determine whether there are two of them that are equal in modulus to $m_F(z')$. If so, z' is in $D(F)$.
(iv) For each dominant root z' construct a *search circle* with centre z' and radius ϵ. Locate, if possible, points z'' on the search circle that also have the dominance and equimodularity properties. Each such point determines an arc $z'z''$ forming part of $D(F)$.
(v) Repeat the local search for each point z'' found in step (iv), thus extending the arcs. We shall refer to this process as *extension*.

The local search can be assisted by using results of Salas and Sokal [10, Section 4.2] concerning the slope of an equimodular curve. But there can be complications caused by singularities.

Example Let $B(z)$ be the 3×3 matrix displayed in Section 6 of [BC], where it is called $T(z)$. The coefficients $b_1(z), b_2(z), b_3(z)$ of the characteristic polynomial are given in the Appendix, and $v(t,z)$ can be calculated by substituting them in the generic polynomial $r_3(t)$ [BC, Section 4]. After removing the factors $z^2(z+1)^2$, we find the special points listed in [BC]. The dominant ones are:

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

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, and (necessarily) all the 0-points are double roots. Thus $-1.8726 + 1.1275i$ is presumably the end-point of a dominant equimodular curve. Using the extension process, this curve is found to pass smoothly through the all the special points with positive imaginary parts. However, complications arise as the curve approaches the point 0.3369 on the real axis, when the search circle will contain three possible points z''. This is because the point 0.3369 is a singularity, a fact that can be verified by the vanishing of the Jacobian (in this case, the discriminant of v considered as a polynomial in z). Continuing the process in the obvious way, we obtain the 'largest' of the three curves depicted in Figure 1 of [BC]. It has four segments, two of which are non-differentiable at the singularity. This curve, $D(B)$, is shown in Figure 4 (Section 7) of this paper. □

We now turn to the case where $F(z)$ is reducible, with just two distinct irreducible constituents, $U(z)$ and $W(z)$. We shall say that a point z^* is *dominant for the pair* (U, W) if $m_U(z^*) = m_W(z^*)$, and denote the set of points dominant for (U, W) by $D(U, W)$. Clearly, $D(U, W)$ is a subset of $E(U, W)$ and

$$D(U, W) \subseteq D(F) \subseteq D(U) \cup D(W) \cup D(U, W).$$

According to Theorem 1, the curves comprising $E(U, W)$ are defined by the vanishing of a polynomial function $v(t, z)$ and the curves can be decomposed into segments, with the additional property that the special points defined by $v(4, z) = 0$ occur in pairs, so that the segments link up to form closed curves. In this case $v(t, z) = q(s, z)\tilde{q}(s, z)$, so all the relevant information can be obtained from the polynomial q. Noting that the values $t = 4$ and $t = 0$ correspond to $s = 1$ and $s = -1$ respectively, we can adapt the method suggested above to the determination of $D(U, W)$.

(i) Substitute the functional forms of the coefficients in the generic polynomial $q(s)$, obtaining a polynomial $q(s, z)$.
(ii) Compute the special points, that is, the roots of $q(1, z) = 0$ and $q(-1, z) = 0$.
(iii) For each root z', calculate $m_U(z')$ and $m_W(z')$, and determine whether they are equal. If so, z' is in $D(U, W)$.
(iv) For each dominant root z' construct a search circle and locate points on it that also have the dominance and equimodularity properties. Each such point determines an arc $z'z''$ forming part of $D(U, W)$.
(v) Repeat the local search for each point z'' found in step (iv), thus extending the arcs.

6. Triple points

An equimodular curve is smooth, except at points where a Jacobian vanishes [BC, Section 6]. 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 z_0 is a *triple point* if three (or more) eigenvalues have equal modulus at z_0. (Salas and Sokal [10] 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 extension process outlined in the previous section will recognise a triple point. As we approach a triple point, extra points will appear on the search circle, 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 $D(F)$. But in the extension process we are constructing an equimodular curve that forms part of $D(F)$. In this case, each of the three curves passing through a triple point z_0 has the property that its points on one side of z_0 are in $D(F)$, while those on the other side are not. See Figure 2.

Figure 2: a typical triple point, bold lines indicating the dominant parts.

Example Let $C(z)$ be the 4×4 matrix whose characteristic polynomial is given in the Appendix. Substituting the coefficients of the characteristic polynomial in the generic polynomial $r_4(t)$ gives the polynomial $v(t, z)$, and after removing the factor $(z+1)^2$ we obtain a polynomial which has degree 6 in t and 16 in z. Only a few of the special points have the dominance property:

$$\begin{aligned}(t=4): &\quad -1.5684 \pm 2.1597i, \quad 0.5000;\\ (t=0): &\quad 0.5324 \pm 1.5856i.\end{aligned}$$

The dominant 4-point at 0.5000 is a double root.

We begin the extension process at one of the 0-points, say $0.5324 + 1.5856i$, and denote by Γ_1 the equimodular curve so constructed. Moving to the left, the extension proceeds smoothly and the curve terminates at the end-point $-1.5684 + 2.1597i$. This part of Γ_1 is dominant.

However, on leaving $0.5324 + 1.5856i$ in the other direction, the extension process hits a triple point α at $0.5872 + 1.4516i$. Two other equimodular curves Γ_2 and Γ_3 pass through α, and they intersect Γ_1 again at $\beta = 0.5944 + 1.2671i$. 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. See Figure 3.

Continuing the extension process along Γ_1, the curve crosses the real axis at the double end point 0.5000. It then traces out a path conjugate to the one already described, including the brief loss of dominance, before terminating at $-1.5684 - 2.1597i$. The complete dominant set for C is shown in Figure 3. It contains parts of three equimodular curves $\Gamma_1, \Gamma_2, \Gamma_3$. The decomposition of Γ_1 into segments is clear from the description given above, but further calculations are needed to determine the decompositions of Γ_2 and Γ_3. □

Figure 3: $D(C)$, with schematic diagram of one pair of triple points.

7. Triple points in the reducible case

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

The application of this criterion may involve the determination of triple points, as indicated in the following theorem. It implies that, in the construction of $D(U,W)$ by the extension process, triple points will occur if and only if the curve under construction hits $D(U)$ or $D(W)$, and such a triple point will separate a part of $D(U)$ or $D(W)$ that belongs to $D(F)$ from a part that does not.

Theorem 2 A triple point that belongs to $D(U,W)$ must belong to $D(U) \cup D(W)$.

Proof Consider part of an equimodular curve that belongs to $D(U,W)$. 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 $|\lambda_1(z)| = m_U(z)$ and $|\mu_1(z)| = m_W(z)$. Then at a triple point z_0 there is a third eigenvalue equal in modulus to $\lambda_1(z_0)$ and $\mu_1(z_0)$. Without loss of generality, we may take it to be an eigenvalue $\mu_2(z_0)$ of $W(z_0)$. It follows that the two other equimodular curves passing through z_0 are defined by the equations

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

Since $|\mu_1(z)| = m_W(z)$ in a neighbourhood of z_0, it follows that the second curve is in $D(W)$, and hence z_0 is in $D(W)$. □

Example Let the matrix $F(z)$ be defined by

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

where $B(z)$ and $C(z)$ are the square matrices of size 3 and 4 whose dominant sets $D(B)$ and $D(C)$ have been described in the examples above. These sets are shown in Figure 4; they intersect at the real point 0.5000.

Figure 4: $D(B)$ (continuous lines) and $D(C)$ (broken lines).

It remains to find $D(B, C)$, and then to determine which parts of $D(B)$ and $D(C)$ also belong to $D(F)$. The calculation of the special points is described in the Appendix, and 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. They are:

$$\begin{aligned} (s = 1): &\quad -1.0000, \quad -0.4660 \pm 1.4456i, \quad 0.6383; \\ (s = -1): &\quad 0.2574 \pm 0.6675i. \end{aligned}$$

When the extension process is used to construct $D(B, C)$, four triple points are encountered. Starting from 0.6383, a dominant equimodular curve Δ_1 extends to the left until it hits triple points on $D(C)$ at $0.5043 \pm 0.1927i$. Here dominance is acquired by another equimodular curve Δ_2 that passes through the dominant special points at $0.2574 \pm 0.6675i$ and $-0.4660 \pm 1.4456i$ before hitting $D(B)$ at the triple points $-0.6735 \pm 1.5822i$. At these points dominance is acquired by another equimodular curve Δ_3 that closes up at the dominant special point -1.0000.

The preceding description also determines the parts of $D(B)$ and $D(C)$ that belong to $D(F)$. 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.

The set $D(F)$ is shown in Figure 5. The subset $D(B, C)$ is the union of parts of three closed curves Δ_1, Δ_2, Δ_3, belonging to $E(B, C)$, and the rest comprises parts of $D(B)$ and $D(C)$.

Figure 5: the dominant set $D(F)$.

8. Conclusion

The matrices $B(z)$ and $C(z)$ discussed in the preceding sections are not totally arbitrary, as they arise in the discussion of the limit points of the chromatic roots of a family of graphs [4,SC]. Nevertheless, their individual properties and their interaction as constituents of $F(z)$ provide good illustrations of some of the difficulties involved in the determination of dominant equimodular curves.

Among these difficulties are the points of non-differentiability, indicated by the vanishing of a Jacobian polynomial. The degree of this polynomial may be large [BC, Section 6]. Another difficulty is the presence of triple points, which are not easy to detect analytically, and which may be confusingly close together. In addition, there is the inherent problem that the end-points of segments are points where two eigenvalues are equal: this means that although the points themselves may be found relatively simply, the behaviour of curves that pass close to them is difficult to determine.

On the positive side, the technique of reducing F to its constituents, and examining them in pairs has some advantages. There are some simple rules governing the topology of $D(F)$, and these can be used in the construction of the required configuration. Further checks are provided by the decomposition of the curves into segments.

References

[BC]. N.L. Biggs. Equimodular curves. *CDAM Research Report Series*, LSE-CDAM-2000-17, September 2000.

[SC]. S-C. Chang. Chromatic polynomials for lattice strips with cyclic boundary conditions. Preprint, October 2000.

The full list of references from [BC] is repeated here for convenience.

1. S. Beraha, J. Kahane, N.J. Weiss. Limits of zeros of recursively defined families of polynomials. *Studies in Foundations and Combinatorics, Advances in Math., Supplementary Studies* 1 (1978) 213-232.

2. N.L. Biggs. A matrix method for chromatic polynomials -II. *CDAM Research Report Series*, LSE-CDAM-2000-4, April 2000.
3. N.L. Biggs, R.M. Damerell, A.D. Sands. Recursive families of graphs. *J. Combinatorial 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, R. Shrock. Ground state entropy of the Potts antiferromagnet on strips of the square lattice. http://uk.arXiv.org/ps/cond-mat/0004161
6. D. Cox, J. Little, D. O'Shea. *Using Algebraic Geometry.* Springer, New York, 1998.
7. 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.
8. E.H. Lieb, A.D. Sokal. A general Lee-Yang theorem for one-component and multicomponent ferromagnets. *Commun. Math. Phys.* 80 (1981) 153-179.
9. R. Read, G.F. Royle. Chromatic roots of families of graphs. In: *Graph Theory, Combinatorics and Applications* (ed. Y.Alavi, G. Chartrand, O.R. Ollerman, A.J.Schwenk) Wiley 1991.
10. J. Salas, A.D. Sokal. Transfer matrices and partition-function zeros for antiferromagnetic Potts models. http://uk.arXiv.org/ps/cond-mat/0004330
11. 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.
12. R. Shrock, S-H. Tsai. Ground-state degeneracy of Potts antiferromagnets: homeomorphism classes with non-compact W boundaries. *Physica A* 265 (1999) 186-223.

Appendix

The purpose of this Appendix is to enable the reader to make an independent check of the calculations referred to in the main part of the paper.

The following are the coefficients of the characteristic polynomials of the matrices $B(z)$ and $C(z)$ discussed in Sections 5,6,7.

$$\begin{aligned} b_1(z) &= -z^4 - 2z^3 - 4z^2 - 1 \\ b_2(z) &= z(z+1)(z^4 + z^3 + 2z^2 + 2) \\ b_3(z) &= -z^2(z+1)^2; \end{aligned}$$

$$\begin{aligned} c_1(z) &= -z^2 + 2z - 2 \\ c_2(z) &= -2z^3 + z^2 - 2z - 1 \\ c_3(z) &= -z^4 + 1 \\ c_4(z) &= (z+1)^2. \end{aligned}$$

We describe briefly the calculations required to determine $D(F)$, where $F(z)$ is the matrix with constituents $B(z)$ and $C(z)$. The equimodular curves $E(B,C)$ are derived from the generic polynomial $r_{3,4}(t) = s^{-12}q(s)\tilde{q}(s)$, where

$$q(s) = \det \begin{pmatrix} s^3 & s^2 b_1 & s b_2 & b_3 & 0 & 0 & 0 \\ 0 & s^3 & s^2 b_1 & s b_2 & b_3 & 0 & 0 \\ 0 & 0 & s^3 & s^2 b_1 & s b_2 & b_3 & 0 \\ 0 & 0 & 0 & s^3 & s^2 b_1 & s b_2 & b_3 \\ 1 & c_1 & c_2 & c_3 & c_4 & 0 & 0 \\ 0 & 1 & c_1 & c_2 & c_3 & c_4 & 0 \\ 0 & 0 & 1 & c_1 & c_2 & c_3 & c_4 \end{pmatrix}.$$

This can be written as a polynomial of degree 12 in s:

$$q(s) = \sum_{i=0}^{12} q_i s^{12-i}.$$

The coefficient q_i is an integral linear combination of terms of the form $\beta\gamma$, where each β is a monomial of weight i in the b's and each γ is a monomial of weight $12-i$ in the c's. For example,

$$q_0 = c_4^3, \quad q_1 = -b_1 c_3 c_4^2, \quad q_2 = b_2 c_3^2 c_4 + b_1^2 c_2 c_4^2 - 2 b_2 c_2 c_4^2.$$

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

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

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

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

The points listed in Section 7 can now be found by solving the equations $q_0(1,z)=0$ and $q_0(-1,z)=0$ and testing the roots for dominance. In particular the point 0.6383 is the largest real root of $p_9(z)=0$. This corresponds to the critical value 2.6383 found by Chang [SC], a value which has special significance in the physical context.