COURANT'S NODAL LINE THEOREM AND ITS DISCRETE COUNTERPARTS

by G. M. L. GLADWELL

(*Department of Civil Engineering, University of Waterloo, Waterloo, Ontario, Canada* N2L 3G1)

and H. ZHU

(*Department of Applied Mathematics, University of Waterloo, Waterloo, Ontario, Canada* N2L 3G1)

[Received 24 July 2000. Revise 12 January 2001]

Summary

Courant's nodal line theorem (CNLT) states that if the Dirichlet eigenvalues of the Helmholtz equation $\Delta u + \lambda\rho u = 0$ for $D \in \mathbb{R}^m$ are ordered increasingly, the nodal set of the nth eigenfunction u_n, which consists of hypersurfaces of dimension $m-1$, divides D into *no more than n sign domains* in which u_n has one sign. We formulate and prove a discrete CNLT for a piecewise linear finite element discretization on a triangular/tetrahedral mesh.

1. Introduction

Courant's nodal line theorem (CNLT) (**1**) is a theorem of wide applicability with a remarkably simple proof based on the minmax property of the Rayleigh quotient. It relates to the Dirichlet eigenfunctions $u(\mathbf{x})$ of elliptic equations, the simplest and most important of which is the Helmholtz equation

$$\Delta u + \lambda\rho u = 0, \qquad \mathbf{x} \in D. \tag{1}$$

Here Δu is the Laplacian, $\rho(\mathbf{x})$ is positive and bounded, and D is a domain in $\mathbb{R}^m$. Equation (1) governs the spatial eigenmodes of a vibrating membrane in $\mathbb{R}^2$, and acoustic standing waves in $\mathbb{R}^3$.

The *nodal set* of $u(\mathbf{x})$ is defined as the set of points $\mathbf{x}$ such that $u(\mathbf{x}) = 0$. It is known (**2**) that for $D \subset \mathbb{R}^m$, the nodal set of an eigenfunction of (1) is locally composed of hypersurfaces of dimension $m-1$. These hypersurfaces cannot end in the interior of D, which implies that they are either closed, or begin and end at the boundary. In particular therefore, in the plane ($m = 2$), the nodal set of an eigenfunction $u(\mathbf{x})$ of (1) is made up of continuous curves, called *nodal lines*, which are either closed, or begin and end on the boundary.

CNLT states that if the eigenvalues $(\lambda_i)_1^n$ of (1) are ordered increasingly, then each eigenfunction $u_n(\mathbf{x})$ corresponding to λ_n, divides D by its nodal set, into at most n subdomains, called *nodal domains*, or the more informative *sign domains* in which $u_n(\mathbf{x})$ has one sign. If λ_n is simple, so that $\lambda_{n-1} < \lambda_n < \lambda_{n+1}$, then the proof of CNLT follows as a straightforward application of Courant's minmax theorem for the Rayleigh quotient. If λ_n is an eigenvalue of multiplicity r, then the straightforward argument used for a simple eigenvalue shows merely that any eigenfunction corresponding to λ_n has at most $n+r-1$ sign domains. Courant and Hilbert (**1**) give a complicated refinement of their argument which reduces the number to n. Herrmann (**3**) and Pleijel (**4**) give a simpler argument which arrives at the same conclusion; this is repeated here in section 2.

Nowadays it is routine to construct approximate solutions of (1) by using some finite element method (FEM). In the simplest implementation for $D \subset \mathbb{R}^2$, D is triangulated and u is taken to be linear in each triangle; for $D \subset \mathbb{R}^3$, u is taken to be linear in each tetrahedron. We are not concerned with whether the approximate solution converges to the actual continuous solution. Instead, we study the FEM solution as a phenomenon in its own right. For a triangular (tetrahedral) mesh in $\mathbb{R}^2$ ($\mathbb{R}^3$), it is a piecewise linear field satisfying the boundary condition $u = 0$ on the piecewise straight (plane) boundary. We ask whether an approximate FEM solution, corresponding to some refined or crude mesh, has any properties analogous to those which the actual solution of (1) possesses. In particular we ask whether FEM solutions satisfy a discrete counterpart of CNLT.

CNLT relates to *signs*, the signs of eigenfunctions. The FEM reduces (1) to a generalized eigenvalue problem of the form

$$(\mathbf{K} - \lambda\mathbf{M})\mathbf{u} = \mathbf{0}. \tag{2}$$

The matrices $\mathbf{K}$, $\mathbf{M}$ are the *stiffness* and *mass matrices* respectively; they are both symmetric and positive definite. For a triangular mesh in $\mathbb{R}^2$ or a tetrahedral mesh in $\mathbb{R}^3$, it is easy to show that $\mathbf{M}$ is non-negative, that is, $m_{ij} \geqslant 0$ for all $i, j = 1, \ldots, n$; its diagonal entries are positive, and the off-diagonal entry m_{ij} is positive if and only if the line connecting vertices P_i and P_j is an edge of the mesh; we write this as $P_i \sim P_j$. Since $\mathbf{K}$ is positive-definite, its diagonal entries k_{ii} are positive; the signs of its non-zero off-diagonal entries depend on characteristics of the mesh. In the simplest case, a triangular mesh in $\mathbb{R}^2$, the non-zero entries are non-positive if (but not only if) all the triangles are acute-angled. If some triangles are obtuse then some off-diagonal entries might be positive. There are similar results for tetrahedral meshes in $\mathbb{R}^3$.

To formulate a discrete counterpart of CNLT for FEM solutions we need to construct analogues of the domains which appear in the continuous case. The analogues are expressed by introducing the concept of a *sign graph*. We show that the eigenvectors obey a discrete counterpart of CNLT. As with the continuous version, the theorem is easily proved for distinct eigenvalues; multiple eigenvalues require delicate treatment.

It is easy to construct counterexamples of meshes with some *obtuse* angled triangles for which the discrete CNLT fails, and others for which it holds.

2. Courant's nodal line theorem

We recall a statement and proof of two versions of CNLT so that we can indicate later how the continuous and discrete results differ from each other. We express the problem in variational form. Define

$$(u, v)_D = \int_D \nabla u \cdot \nabla v \, d\mathbf{x}, \quad \langle u, v\rangle_D = \int_D \rho u v \, d\mathbf{x}. \tag{3}$$

It is well known that (1) has infinitely many positive eigenvalues

$$0 < \lambda_1 \leqslant \lambda_2 \leqslant \lambda_3 \leqslant \cdots. \tag{4}$$

The corresponding eigenfunctions $\{u_i\}_1^\infty$, where the u_i are orthonormal, form a complete set of $L^2(D)$. We say that $u \in H_0^1(D)$ is a weak solution of (1) if

$$-(u, v)_D + \lambda\langle u, v\rangle_D = 0 \tag{5}$$

for all $v \in H_0^1(D)$. The fundamental theorem for the Rayleigh quotient

$$\lambda_R \equiv \frac{(u, u)_D}{\langle u, u\rangle_D} \geqslant \lambda_n \tag{6}$$

is that if $\langle u, u_i\rangle_D = 0, i = 1, 2, \ldots, n-1$, then $\lambda_R \geqslant \lambda_n$, with equality if and only if $u(\mathbf{x}) = u_n(\mathbf{x})$.

THEOREM 1 (Courant–Hilbert). *Suppose the eigenvalues λ_i of* (1) *are ordered increasingly, as in* (4)*, and $u_n(\mathbf{x})$ is an eigenfunction corresponding to λ_n. If λ_n has multiplicity $r \geqslant 1$, so that*

$$\lambda_{n-1} < \lambda_n = \lambda_{n+1} = \cdots = \lambda_{n+r-1} < \lambda_{n+r}, \tag{7}$$

$u_n(\mathbf{x})$ has at most $n + r - 1$ sign domains.

Proof. Suppose $u_n(\mathbf{x})$ has m sign domains D_i such that $\bigcup_{i=1}^m D_i = D$. Define

$$w_i(\mathbf{x}) = \begin{cases} \beta_i u_n(\mathbf{x}), & \mathbf{x} \in D_i, \\ 0, & \text{otherwise}, \end{cases} \tag{8}$$

and take

$$v(\mathbf{x}) = \sum_{i=1}^m c_i w_i(\mathbf{x}), \quad \sum_{i=1}^m c_i^2 = 1. \tag{9}$$

Since the D_i are disjoint, $(w_i(\mathbf{x}))_1^m$ are orthogonal. Scale the w_i, that is, choose the β_i, so that $\langle w_i, w_i\rangle_{D_i} = 1$, then

$$\langle v, v\rangle = \sum_{i=1}^m c_i^2\langle w_i, w_i\rangle_{D_i} = \sum_{i=1}^m c_i^2 = 1.$$

Since $w_i(\mathbf{x})$ satisfies (5) with $\lambda = \lambda_n$ in D_i, and $w_i(\mathbf{x}) = 0$ on ∂D_i,

$$(v, v) = \sum_{i=1}^m c_i^2(w_i, w_i)_{D_i} = \sum_{i=1}^m c_i^2\lambda_n\langle w_i, w_i\rangle_{D_i} = \lambda_n.$$

Thus $\lambda_R = \lambda_n$. But we may choose $(c_i)_1^m$ so that $\langle v, u_i\rangle = 0, i = 1, 2, \ldots, m-1$, and hence, for that choice, $\lambda_R \geqslant \lambda_m$. Thus $\lambda_m \leqslant \lambda_n$. Since $\lambda_n < \lambda_{n+r}$, we have $\lambda_m < \lambda_{n+r}$ so that $m < n + r$; $m \leqslant n + r - 1$.

Note that this proof does not require that D be connected. To reduce the upper bound on m from $n + r - 1$ to n, we need the unique continuation theorem (**5**): *If any solution $u \in H_0^1(D)$ of* (1) *vanishes on a non-empty open subset of a connected domain D, then $u \equiv 0$ in D.* We can then conclude as follows.

THEOREM 2 (Herrmann–Pleijel). *Suppose D is connected, the eigenvalues are ordered increasingly,* $0 < \lambda_1 < \lambda_2 \leqslant \cdots$, *and $u_n(\mathbf{x})$ is an eigenfunction corresponding to λ_n; then $u_n(\mathbf{x})$ has at most n sign domains.* (*Recall that when D is connected, λ_1 is simple*: $0 < \lambda_1 < \lambda_2$ (**6**)).

Proof. Suppose $u_n(\mathbf{x})$ has $m > n$ sign domains. Define the $w_i(\mathbf{x})$ as before, and define $v(\mathbf{x})$ by (9) with $c_{n+1} = 0 = \cdots = c_m$, so that $v(\mathbf{x}) \equiv 0$ on $D_{n+1}, \ldots, D_m$. Again we have $\lambda_R = \lambda_n$, and we may choose $(c_i)_1^n$ so that $\langle v, u_i \rangle = 0$, $i = 1, 2, \ldots, n-1$. Thus $v(\mathbf{x}) \in H_0^1(D)$ is an eigenfunction of (1), but it is identically zero on D_{n+1} and hence, by the unique continuation theorem, it is identically zero on D. This contradiction implies $m \leqslant n$.

COROLLARY 1 (Courant–Hilbert). *Theorem* 2 *holds even if D is not connected.*

Proof. Suppose D consists of p connected domains $(D_k)_1^p$. Label the eigenvalues $\lambda_i^{(k)}$ of each D_k increasingly and suppose the corresponding eigenfunctions are $u_i^{(k)}(\mathbf{x})$. Now assemble the eigenvalue sequences $\{\lambda_i^{(k)}\}$, $k = 1, 2, \ldots, p$; $i = 1, 2, \ldots$ into one non-decreasing sequence $\{\lambda_j\}$ to give the eigenvalues of D. The corresponding eigenfunctions of D are

$$u_j(\mathbf{x}) = \begin{cases} u_i^{(k)}(\mathbf{x}) & \text{on } D_k, \\ 0 & \text{elsewhere.} \end{cases}$$

The ordinal number j of a given $\lambda_i^{(k)}$ in this sequence will satisfy $j \geqslant i$. Theorem 2 for D_k states that $u_i^{(k)}(\mathbf{x})$ has no more than i sign domains on D_k, so that $u_j(\mathbf{x})$ will have no more than j sign domains on D_k, and it will be zero elsewhere.

3. Sign characteristics of the element matrices

The sign characteristics of the FEM eigenvectors depend on those of the element stiffness and mass matrices, which we now study.

With the notation of (3), the variational statement for the Rayleigh quotient is

$$\lambda_i = \min_{\substack{\langle u, u_j \rangle = 0 \\ j = 1, 2, \ldots, i-1}} \frac{(u, u)}{\langle u, u \rangle} = \frac{(u_i, u_i)}{\langle u_i, u_i \rangle}. \tag{10}$$

This leads to the matrix eigenvalue equation (2) where, following the usual FEM procedure,

$$\int_D (\nabla u)^2 \, d\mathbf{x} \backsimeq \mathbf{u}^T \mathbf{K} \mathbf{u}, \quad \int_D \rho u^2 \, d\mathbf{x} \backsimeq \mathbf{u}^T \mathbf{M} \mathbf{u}; \tag{11}$$

the symbol $\backsimeq$ indicates that, in applying the FEM we replace the (curved) boundary of D by an appropriate piecewise-linear boundary; the discrete solution satisfies $u = 0$ on this piecewise-linear boundary.

The global matrices $\mathbf{K}$, $\mathbf{M}$ are obtained by assembling the element matrices $\mathbf{K}_e$, $\mathbf{M}_e$. It may be shown (**7**) that for a triangular element in $\mathbb{R}^2$, all the elements in $\mathbf{M}_e$ are strictly positive, that is, $\mathbf{M}_e > 0$. If the triangle is acute, then all the off-diagonal terms in $\mathbf{K}_e$ are negative: $\mathbf{K}_e$ has the sign pattern

$$\mathbf{K}_e = \begin{bmatrix} + & - & - \\ - & + & - \\ - & - & + \end{bmatrix}.$$

Although, as we have already stated, we are not primarily interested in the *accuracy* of the FEM solutions, we note that, for accuracy, meshes are judged according to their *ratio of mesh quality*. For a triangular mesh, this ratio is optimum for an equilateral triangle and worsens as a triangle becomes obtuse.

For a tetrahedral element with linear interpolation, $\mathbf{M}_e$ will be positive, while $\mathbf{K}_e$ will have the sign pattern

$$\mathbf{K}_e = \begin{bmatrix} + & - & - & - \\ - & + & - & - \\ - & - & + & - \\ - & - & - & + \end{bmatrix}$$

provided that the angles between the outward drawn normals to the faces are all *obtuse*.

We conclude that if the finite element mesh is made up of elements satisfying the stated constraints then the mesh is an undirected graph $\mathcal{G}$ on the N vertices P_i, with the following properties, referred to as Conditions 1 to 3:

1. **K** is symmetric, positive semi-definite with non-positive off-diagonal terms; $k_{ij} < 0$ if and only if vertices P_i, P_j are the ends of an edge; in this case we say vertices P_i, P_j are *adjacent*, and write as $P_i \sim P_j$;
2. **M** is symmetric, positive definite, with non-negative off-diagonal terms; *and* **M** has the same pattern of zeros as **K**, that is, if $i \neq j$, then $m_{ij} > 0$ if and only if $k_{ij} < 0$;
3. For an FEM mesh on a *connected* domain D, the mesh forms a *connected graph* $\mathcal{G}$. It is known (**8**) that the matrices **K** and **M** are then *irreducible*.

4. Properties of FEM eigenvectors

As we noted in section 2, one of the mainstays of theory related to (1) is the unique continuation theorem (**5**): if an eigenfunction $u(\mathbf{x})$ is zero in a finite region of a connected domain D, then it is identically zero. There does not appear to be a straightforward analogue of this result for FEM eigenvectors: an eigenvector can be zero in one or more complete elements without being identically zero, as the example in Fig. 1 shows.

There are, however, some simple properties of eigenvectors. To obtain these we first note that vertices of an FEM mesh are of three kinds: *boundary* vertices, where $u \equiv 0$ because of the boundary conditions; vertices adjacent to boundary vertices, which we call *near-boundary* vertices; the remainder, which we term *interior* vertices.

THEOREM 3. *Under Conditions* 1 *and* 2 *a nodal vertex of an eigenvector which is not a boundary vertex cannot be isolated.*

Proof. If $u_i = 0$, the ith line of (2) is

$$\sum (k_{ij} - \lambda m_{ij}) u_j = 0, \tag{12}$$

where the sum is over those j ($\neq i$) for which $P_i \sim P_j$; for those j, $k_{ij} - \lambda m_{ij} < 0$. If $u_j \geqslant 0$ ($\leqslant 0$) for all such j, with one inequality strict, then the left-hand side of (12) is strictly negative (positive), which is a contradiction.

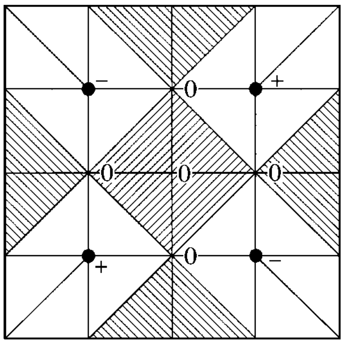

Fig. 1 An eigenvector can have zero (shaded) polygons

This implies that each vertex in a nodal set must *either* have both positive and negative neighbours, *or* all nodal neighbours: a nodal set separates positive and negative vertex sets.

There is an important *maximum principle* for (1): a solution $u(\mathbf{x})$ cannot have an interior positive minimum or an interior negative maximum (**9**). The discrete counterpart is given next.

THEOREM 4. *An eigenvector of* (2) *cannot have a positive minimum or a negative maximum at an interior vertex.*

Proof. Because of the way it is defined, by (11), the stiffness matrix $\mathbf{K}_e$ of an interior element admits a rigid body mode: $\mathbf{K}_e\{1, 1, 1\} = \mathbf{0}$. If P_i is an interior vertex, all the elements to which P_i belongs are interior elements. This means that after assembling the $\mathbf{K}_e$ to form $\mathbf{K}$ we may deduce that, if P_i is an interior vertex, then $\sum_j k_{ij} = 0$. The ith line of (2) is

$$\begin{aligned} 0 &= \sum_j k_{ij}u_j - \lambda \sum_j m_{ij}u_j \\ &= \sum_j k_{ij}(u_j - u_i) + \left(\sum_j k_{ij}\right) u_i - \lambda \sum_j m_{ij}u_j. \end{aligned} \tag{13}$$

Suppose that there is a local positive minimum at an interior vertex P_i, so that $u_i \geqslant 0$ and $u_j - u_i \geqslant 0$, and either the first inequality is strict, or the second inequality is strict for at least one j such that $P_j \sim P_i$. The second sum in (13) is zero, while the other two are both non-positive, with at least one being *negative*. This is impossible.

In order to formulate the FEM counterpart of CNLT we must define an FEM replacement for the *sign domains* of the continuous theorem. For simplicity consider the (acute-angled) triangular mesh over a domain $D \subset \mathbb{R}^2$ with $u = 0$ on the piecewise straight boundary ∂D. There are two distinct ways of looking at the piecewise linear function u obtained from an eigenvector of (2): by looking at the values u_i, and particularly the signs of u_i, at the vertices P_i of the FEM mesh; by looking at the subregions with piecewise straight boundaries on which the linearly interpolated $u(x, y)$ has one sign, either loosely $u(x, y) \geqslant (\leqslant)\ 0$ or strictly $u(x, y) > (<)\ 0$.

Consider the first way. The FEM mesh defines a graph $\mathcal{G}$ with N vertices P_i. An FEM vector

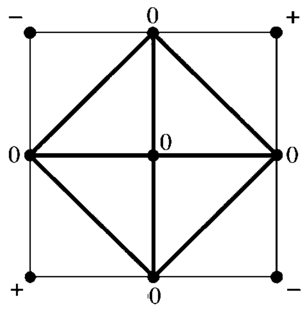

Fig. 2 Positive, negative and zero sign graphs

$\mathbf{u} \in \mathbb{R}^N$ associates a value u_i, and in particular a sign $+$, $-$, or 0, to each vertex $P_i \in \mathcal{G}$ (Fig. 2). We may connect the positive vertices by edges to form *maximal connected subgraphs* of $\mathcal{G}$, called *positive sign graphs*, denoted by S. We may do the same with the negative vertices. In this way we partition $\mathcal{G}$ into disjoint positive and negative sign graphs, and zero vertices. We use the notation SG($\mathbf{u}$) to denote the number of sign graphs of $\mathbf{u}$. Two sign graphs S_1, S_2 are said to be *adjacent* if there are vertices $P_1 \in S_1$, $P_2 \in S_2$ such that $P_1 \sim P_2$. We need the following simple but important property.

LEMMA 1. *If two different sign graphs are adjacent they have opposite signs.*

Proof. If they had the same sign then neither would be maximal.

Now consider the second way: looking at the sign of the piecewise linear 'eigenfunction' interpolated from the vertex values u_i (of an eigenvector $\mathbf{u}$). The domain D may be divided into subdomains D_i in each of which u is positive, negative or identically zero, and on the boundaries of which, $u = 0$. Each of these subdomains will be polygonal in $\mathbb{R}^2$ (polyhedral in $\mathbb{R}^3$). In particular the *nodal places* of u in $\mathbb{R}^2$ will be piecewise straight lines, either closed or beginning and ending on the boundary, or nodal polygons. In $\mathbb{R}^3$ they are piecewise plane surfaces which are either closed or begin and end on the boundary, or polyhedra.

For triangular or tetrahedral meshes satisfying Conditions 1 and 2, there is a clear correspondence between the sign graphs on the one hand and the sign domains on the other. Inside each positive, negative or zero sign domain there is respectively just one positive, negative or zero sign graph. This means that we can count the sign domains by counting sign graphs. In the remainder of the paper we consider only sign *graphs*.

We note, however, that the rectangular FEM which is sometimes used in $\mathbb{R}^2$ does not have such simple properties. Inside a rectangle, $u(x, y)$ has a bilinear interpolation

$$u(x, y) = p + qx + ry + sxy.$$

Now all four vertices of the rectangle are neighbours of each other, in the sense that all the off-diagonal elements in the element matrices are non-zero. This is why we show the vertices of the rectangle joined by the diagonals as well as by the sides. (But the intersection of the diagonals is not a vertex of the graph.) It may be shown that the element mass matrix is strictly positive, and

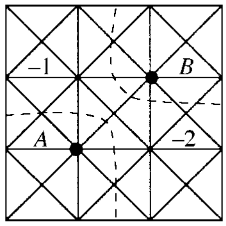

Fig. 3 *A* and *B* have the same sign, are adjacent, but are situated in different sign domains

that the off-diagonal elements of the element stiffness matrix $\mathbf{K}_e$ are strictly negative if and only if the sides a, b of the rectangle satisfy $1/\sqrt{2} < a/b < \sqrt{2}$. There is a similar result for a rectangular box mesh in $\mathbb{R}^3$. Thus, under these conditions, the matrices $\mathbf{K}$, $\mathbf{M}$ for the whole mesh will satisfy Conditions 1 and 2. This means that we can apply the analysis given below to the sign *graphs* of a rectangular mesh. However, for a rectangular mesh there is no longer a one-to-one correspondence between sign *graphs* and sign *domains*. The example in Fig. 3 shows a mesh made up of nine square elements. The vertices *A* and *B* are adjacent and have the same sign, so that they belong to the same sign *graph*. However, because nodal lines in an element are now hyperbolic, not straight, *A* and *B* lie in different sign *domains*; there is an intervening negative sign domain between them.

5. A discrete CNLT

We now state and prove a discrete version of CNLT, for a triangular or tetrahedral mesh.

THEOREM 5. *Suppose Conditions* 1 *and* 2 *are satisfied. If the eigenvalues are labelled increasingly, as in* (4)*, and* λ_n *has multiplicity r, as in* (7)*, then* $\mathrm{SG}(\mathbf{u}_n) \leqslant n + r - 1$.

We need a preliminary result. Suppose $\mathbf{u} \in \mathbb{R}^{N\times 1}$, an arbitrary vector, has m sign graphs $(S_k)_1^m$, where $m \leqslant N$. Label the vertices so that the vertex index i runs through the sign graphs $S_1, S_2, \ldots, S_m$, and then through any zero vertices. Thus partition $\mathbf{u}$ in the form

$$\mathbf{u}^T = \{\mathbf{v}_1^T, \mathbf{v}_2^T, \ldots, \mathbf{v}_m^T, \mathbf{0}\}. \tag{14}$$

Now construct the vectors $\mathbf{w}_i^T = \{\mathbf{0}, \mathbf{0}, \ldots, \mathbf{v}_i^T, \mathbf{0}, \ldots, \mathbf{0}\}$ so that $\mathbf{u} = \sum_{i=1}^m \mathbf{w}_i$, and consider

$$\mathbf{v} = \sum_{i=1}^m c_i \mathbf{w}_i. \tag{15}$$

Using straightforward algebra we may establish the following generalization of the powerful (**10**, Lemma 5).

LEMMA 2.

$$\mathbf{v}^T(\mathbf{K} - \lambda\mathbf{M})\mathbf{v} = \sum_{i=1}^m c_i^2 \mathbf{w}_i^T(\mathbf{K} - \lambda\mathbf{M})\mathbf{u} - \tfrac{1}{2}\sum_{i,j=1}^m a_{ij}(c_i - c_j)^2, \tag{16}$$

where $a_{ij} = \mathbf{w}_i^T(\mathbf{K} - \lambda\mathbf{M})\mathbf{w}_j$.

Proof of Theorem 5. Suppose $\mathbf{u} = \mathbf{u}_n$ has m sign graphs. Since $(\mathbf{K} - \lambda_n \mathbf{M})\mathbf{u} = \mathbf{0}$, Lemma 2 gives

$$\mathbf{v}^T(\mathbf{K} - \lambda_n \mathbf{M})\mathbf{v} = -\tfrac{1}{2} \sum_{i,j=1}^{m} a_{ij}(c_i - c_j)^2. \tag{17}$$

If the sign graphs S_i, S_j are not adjacent, $a_{ij} \equiv 0$. If S_i, S_j are adjacent, $\mathbf{w}_i$ and $\mathbf{w}_j$ have opposite signs. Since $\mathbf{w}_i$ and $\mathbf{w}_j$ are non-overlapping, the products in a_{ij} involve only non-zero off-diagonal entries in $\mathbf{K}$ and $\mathbf{M}$; those in $\mathbf{K}$ are negative, those in $\mathbf{M}$ are positive. Thus $a_{ij} = (+)(-)(-) - (+)(+)(-) = +$. Thus, $\mathbf{v}^T\mathbf{K}\mathbf{v} - \lambda_n \mathbf{v}^T\mathbf{M}\mathbf{v} \leqslant 0$, that is,

$$\lambda_R = \frac{\mathbf{v}^T\mathbf{K}\mathbf{v}}{\mathbf{v}^T\mathbf{M}\mathbf{v}} \leqslant \lambda_n. \tag{18}$$

Let $(\mathbf{u}_i)_1^{m-1}$ be a basis for the eigenspaces of the eigenvalues $(\lambda_i)_1^{m-1}$. Choose the $(c_i)_1^m$ to make $\mathbf{v}$ in (15) $\mathbf{M}$-orthogonal to $(\mathbf{u}_i)_1^{m-1}$, that is, $\mathbf{v}^T\mathbf{M}\mathbf{u}_i = 0$ for $j = 1, 2, \ldots, m-1$. By the minmax theorem for the Rayleigh quotient

$$\lambda_R \geqslant \lambda_m. \tag{19}$$

The inequalities (18) and (19) imply $\lambda_m \leqslant \lambda_n$. Since $\lambda_n < \lambda_{n+r}$ we have $\lambda_m < \lambda_{n+r}$, that is, $m < n + r$ and $m \leqslant n + r - 1$.

6. Further results for multiple eigenvalues

In section 2, we stated two theorems for CNLT related to (1): Theorem 1 showed that if λ_n has multiplicity r, then $u_n(x)$ had at most $n + r - 1$ sign domains. Theorem 1 did not require that D be connected; Theorem 2 reduced the upper bound from $n + r - 1$ to n, and relied on the unique continuation principle for eigenfunctions on a connected D. (Corollary 1 removed the restriction that D be connected.) Theorem 5 is the discrete analogue of Theorem 1: it states $\mathrm{SG}(\mathbf{u}_n) \leqslant n+r-1$. Because there is no discrete analogue of the unique continuation principle, we cannot argue as in Theorem 2 to reduce the bound on $\mathrm{SG}(\mathbf{u}_n)$ from $n+r-1$ to n (when $r > 1$). In fact, it has long been known that an eigenvector of (2) on a connected graph of $\mathcal{G}$ can have *more* than n sign domains. Friedman (**11**) gave the simplest example. If

$$\mathbf{K} = \begin{bmatrix} 4 & -1 & -1 & -1 \\ -1 & 1 & & \\ -1 & & 1 & \\ -1 & & & 1 \end{bmatrix}, \quad \mathbf{M} = \begin{bmatrix} 1 & & & \\ & 1 & & \\ & & 1 & \\ & & & 1 \end{bmatrix},$$

the eigenvalues are $\lambda_1 = 0$, $\lambda_2 = \lambda_3 = 1$, $\lambda_4 = 3$. One eigenvector corresponding to λ_2 is $\{0, 1, 1, -2\}$ as shown in Fig. 4: it has 3 (>2) sign graphs. In (**12**) de Verdière gave a similar example. Duval and Reiner (**10**) attempted to prove $\mathrm{SG}(\mathbf{u}_n) \leqslant n$ even when λ_n was multiple; we

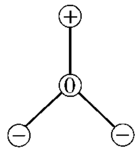

Fig. 4 The eigenvector corresponding to λ_2 has three sign graphs

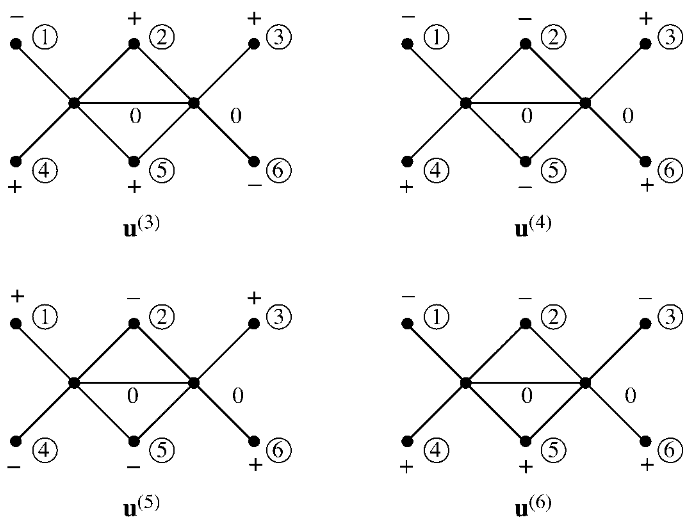

Fig. 5 There are four **M**-orthogonal eigenvectors corresponding to the four-fold eigenvalue $\lambda_3 = \lambda_4 = \lambda_5 = \lambda_6$. Each eigenvector has six sign graphs

discuss the flaws in their logic elsewhere (**7**). We will use the matrix pair

$$\mathbf{K} = \begin{bmatrix} 7 & 0 & 0 & -1 & 0 & 0 & 0 & 0 \\ 0 & 7 & 0 & -1 & -1 & 0 & 0 & 0 \\ 0 & 0 & 7 & 0 & -1 & 0 & 0 & 0 \\ -1 & -1 & 0 & 5 & -3 & -1 & -1 & 0 \\ 0 & -1 & -1 & -3 & 5 & 0 & -1 & -1 \\ 0 & 0 & 0 & -1 & 0 & 7 & 0 & 0 \\ 0 & 0 & 0 & -1 & -1 & 0 & 7 & 0 \\ 0 & 0 & 0 & 0 & -1 & 0 & 0 & 7 \end{bmatrix}, \quad \mathbf{M} = \begin{bmatrix} 3 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 3 & 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 3 & 0 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 5 & 3 & 1 & 1 & 0 \\ 0 & 1 & 1 & 3 & 5 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & 0 & 3 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 & 3 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 3 \end{bmatrix}$$

as our illustrative example. This has a four-fold eigenvalue $\lambda_3 = \lambda_4 = \lambda_5 = \lambda_6$. Each eigenvector in Fig. 5 has $6 = n + r - 1 = 3 + 4 - 1$ sign graphs.

Although each of the eigenvectors shown in Fig. 5 has $\mathrm{SG}(\mathbf{u}) = 6$, there is another **M**-orthogonal basis for the eigenspace, shown in Fig. 6, for which $\mathrm{SG}(\mathbf{u}_j) \leqslant j$ for $j = 3, 4, 5, 6$. The next lemma shows that we can always find such a basis for an eigenspace.

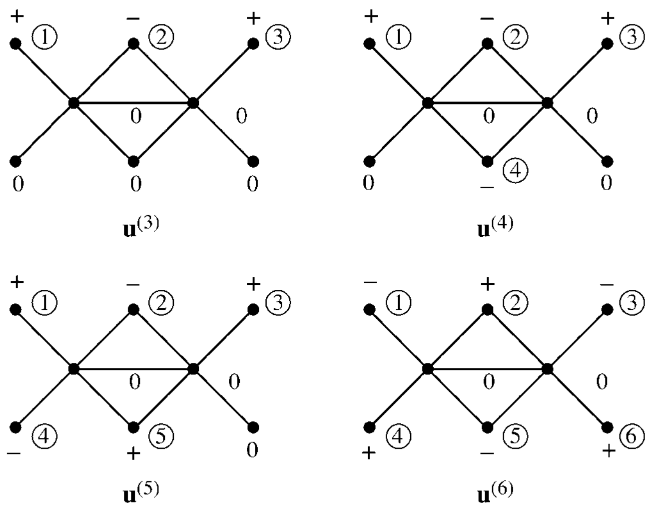

Fig. 6 There are four **M**-orthogonal eigenvectors corresponding to the four-fold eigenvalue $\lambda_3 = \lambda_4 = \lambda_5 = \lambda_6$, such that $\mathrm{SG}(\mathbf{u}_j) \leqslant j$ for $j = 3, 4, 5, 6$

LEMMA 3. *Suppose Conditions* 1 *and* 2 *hold,* $(\mathbf{u}_i)_1^{n-1}$ *is a basis for the eigenspaces of eigenvalues less than* λ_n*, and* $\mathbf{u}$ *is an eigenvector corresponding to* λ_n *such that* $\mathrm{SG}(\mathbf{u}) = m > n$*. Then, in the notation of* (15)*, if* $\mathbf{v}$ *is any non-trivial linear combination of* $(\mathbf{w}_j)_1^n$ *that is* **M***-orthogonal to* $(\mathbf{u}_i)_1^{n-1}$*, then* $\mathbf{v}$ *is an eigenvector corresponding to* λ_n *and* $\mathrm{SG}(\mathbf{v}) \leqslant n$*; furthermore, such a* $\mathbf{v}$ *always exists.*

Proof. We can choose c_j, not all zero, such that $\mathbf{v}$ is **M**-orthogonal to $(\mathbf{u}_i)_1^{n-1}$. By the minmax theorem, $\lambda_R \geqslant \lambda_n$. By Lemma 2, $\lambda_R \leqslant \lambda_n$. Thus $\lambda_R = \lambda_n$ and $\mathbf{v}$ is an eigenvector corresponding to λ_n. By its construction $\mathrm{SG}(\mathbf{v}) \leqslant n$.

We denote an **M**-normalized $\mathbf{v}$ so formed by $\mathbf{v} = T((\mathbf{w}_j)_1^n, (\mathbf{u}_i)_1^{n-1})$. This $\mathbf{v}$ may not be unique; there is always a non-trivial set $(c_j)_1^n$, but it need not be unique.

Note that, in Theorem 2, for the continuous CNLT, we supposed that the eigenfunction $u_n(\mathbf{x})$ had more than n sign domains, and we constructed a purported eigenfunction $v(\mathbf{x})$ orthogonal to $(u_i(\mathbf{x}))_1^{n-1}$, but zero on D_{n+1}; then we used unique continuation of an eigenfunction on a connected D to show that $v(\mathbf{x}) \equiv 0$ in D; this contradicted the hypothesis that $v(\mathbf{x})$ was an eigenfunction, that is, not trivial. In the discrete case we start with an eigenvector $\mathbf{u}_n$ such that $\mathrm{SG}(\mathbf{u}_n) > n$, and construct another eigenvector $\mathbf{v}$ such that $\mathrm{SG}(\mathbf{v}) \leqslant n$; the new eigenvector $\mathbf{v}$ has been formed by deleting one or more of the sign graphs in $\mathbf{u}_n$, but it *is* an eigenvector and there is no contradiction involved.

THEOREM 6. *Suppose Conditions* 1 *and* 2 *hold. If* λ_n *is an eigenvalue of* (2) *of multiplicity r, then we may find r* **M***-orthonormal eigenvectors* $(\mathbf{u}_j)_n^{n+r-1}$ *corresponding to* λ_n *such that* $\mathrm{SG}(\mathbf{u}_j) \leqslant j$*,* $j = n, n+1, \ldots, n+r-1$.

Proof. The r-dimensional eigenspace V of λ_n has an **M**-orthonormal basis $(\mathbf{v}_j)_n^{n+r-1}$. If $\mathrm{SG}(\mathbf{v}_n) \leqslant n$, take $\mathbf{u}_n = \mathbf{v}_n$; otherwise $\mathrm{SG}(\mathbf{v}_n) > n$. In this case if $(\mathbf{w}_j)_1^m$ $(m > n)$ are the sign

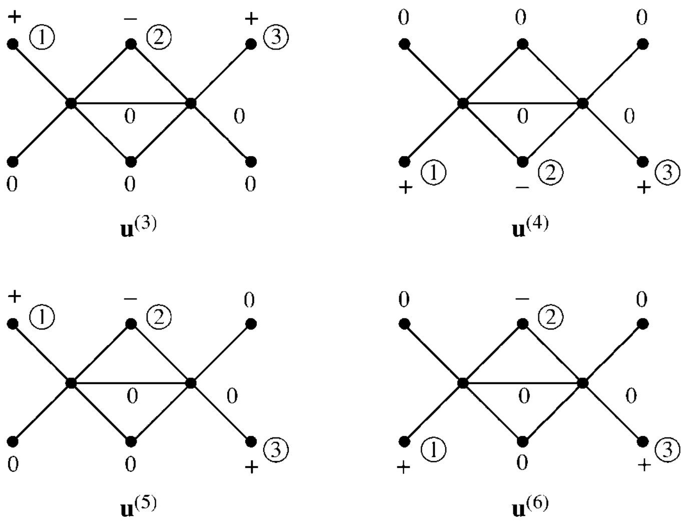

Fig. 7 There are four **M**-orthogonal eigenvectors corresponding to the four-fold eigenvalue $\lambda_3 = \lambda_4 = \lambda_5 = \lambda_6$. Each eigenvector has three sign graphs

graph vectors of $\mathbf{v}_n$, take $\mathbf{u}_n = T((\mathbf{w}_j)_1^n, (\mathbf{u}_i)_1^{n-1})$. We now proceed by induction. Suppose we have constructed **M**-orthonormal $\mathbf{u}_n, \mathbf{u}_{n+1}, \ldots, \mathbf{u}_{n+s-1}$ $(1 < s < r)$ such that $\mathrm{SG}(\mathbf{u}_j) \leqslant j$, $j = n, n+1, \ldots, n+s-1$. We show how to construct $\mathbf{u}_{n+s}$. First find a new orthonormal basis $(\mathbf{u}_j)_1^{n+s-1}$, $(\mathbf{x}_j)_{n+s}^{n+r-1}$ for V. If $\mathrm{SG}(\mathbf{x}_{n+s}) \leqslant n+s$, then take $\mathbf{u}_{n+s} = \mathbf{x}_{n+s}$; otherwise, $\mathrm{SG}(\mathbf{x}_{n+s}) > n+s$; in this case if $(\mathbf{w}_j)_1^m$ $(m > n+s)$ are the sign graph vectors of $\mathbf{x}_{n+s}$, take $\mathbf{u}_{n+s} = T((\mathbf{w}_j)_1^{n+s}, (\mathbf{u}_i)_1^{n+s-1})$. We may proceed in this way to find $(\mathbf{u}_j)_n^{n+r-1}$ such that $\mathrm{SG}(\mathbf{u}_j) \leqslant j$.

Now we note that the eigenspace spanned by the **M**-orthonormal eigenvectors in Fig. 5 or 6 has another basis of not-necessarily **M**-orthogonal eigenvectors shown in Fig. 7; each of these four eigenvectors has 3 (that is, $\leqslant n$) sign graphs. To show that it is always possible to choose such a basis, we prove the next theorem.

THEOREM 7. *Suppose Conditions* 1 *and* 2 *hold, and that* λ_n *is an* r*-fold eigenvalue with eigenspace* V. *There is a basis* $(\mathbf{u}_j)_n^{n+r-1}$ *for* V *such that* $\mathrm{SG}(\mathbf{u}_j) \leqslant n$.

Proof. We proceed much as in Theorem 6. We construct $\mathbf{u}_n$ as before, and then use induction: we suppose that we have found a basis $(\mathbf{u}_j)_n^{n+s-1}$, $(\mathbf{x}_j)_{n+s}^{n+r-1}$ for V such that $\mathrm{SG}(\mathbf{u}_j) \leqslant n$ for $j = n, n+1, \ldots, n+s-1$, and we show how to construct $\mathbf{u}_{n+s}$. If $\mathrm{SG}(\mathbf{x}_{n+s}) \leqslant n$, then $\mathbf{u}_{n+s} = \mathbf{x}_{n+s}$; otherwise $\mathrm{SG}(\mathbf{x}_{n+s}) = n+t$, $1 \leqslant t \leqslant r-1$. In this case, let W be the space spanned by the sign graph vectors $(\mathbf{w}_j)_1^{n+t}$ of $\mathbf{x}_{n+s}$: if $\mathbf{w} \in W$, $\mathbf{w} = \sum_{j=1}^{n+t} c_j \mathbf{w}_j = \mathbf{Wc}$. Let Y be the subspace of W orthogonal to $(\mathbf{u}_j)_1^{n-1}$; Y is not empty because $\mathbf{x}_{n+s} = \sum_{j=1}^{n+t} \mathbf{w}_j \in Y$. If $\mathbf{y} \in Y$, then $\mathbf{y} = \mathbf{Wc}$ and $\mathbf{u}_j^T \mathbf{MWc} = 0$, $j = 1, 2, \ldots, n-1$. Of these $n-1$ constraints on the c_j, say $m \leqslant n-1$ are independent; they may be written as $\mathbf{Bc} = 0$, where $\mathbf{B} = \mathbf{B}(m \times (n+t))$. The matrix $\mathbf{B}$ has

m independent columns which, by suitably numbering the $\mathbf{w}_j$, may be taken as the first m. Thus $\mathbf{Bc} = \mathbf{0}$ may be written as

$$[\mathbf{B}_1, \mathbf{B}_2]\{\mathbf{c}_1, \mathbf{c}_2\} = \mathbf{0}, \tag{20}$$

where $\mathbf{B}_1 = \mathbf{B}_1(m \times m)$ is non-singular, $\mathbf{B}_2 = \mathbf{B}_2(m \times (n + t - m))$,

$$\mathbf{c}_1 = \{c_1, c_2, \ldots, c_m\} \quad \text{and} \quad \mathbf{c}_2 = \{c_{m+1}, \ldots, c_{n+t}\}.$$

The solution space of (20) is spanned by the $n + t - m$ solutions obtained by taking $c_k = \delta_{ik}$, $k = m + 1, \ldots, n + t$, and then solving for $\mathbf{c}_1$. Each such choice gives a vector $\mathbf{y}_i = \mathbf{Wc}^{(i)}$; these vectors are linearly independent because the columns of $\mathbf{W}$, the vectors $\mathbf{w}_j$, are linearly independent and they span Y; by construction $\text{SG}(\mathbf{y}_i) \leqslant m + 1 \leqslant n$. At least one of the $\mathbf{y}_i$, say $\mathbf{y}_p$, must be linearly independent of $(\mathbf{u}_j)_n^{n+s-1}$, for $\mathbf{x}_{n+s} \in Y$ is, by construction, linearly independent of $(\mathbf{u}_j)_n^{n+s-1}$. Take $\mathbf{u}_{n+s} = \mathbf{y}_p$, then $\text{SG}(\mathbf{u}_{n+s}) \leqslant n$. We may proceed in this way to find $(\mathbf{u}_j)_n^{n+r-1}$ such that $\text{SG}(\mathbf{u}_j) \leqslant n$.

We conclude this section by discussing some other implications of Lemma 2.

Suppose that $\mathbf{u}$ is an eigenvector corresponding to a multiple λ_n, so that $\mathbf{Ku} = \lambda_n \mathbf{Mu}$. Suppose that $\text{SG}(\mathbf{u}) = m > n$, and $\mathbf{v}$, given by (15), has been computed so that it is $\mathbf{M}$-orthogonal to $(\mathbf{u}_j)_1^{n-1}$. Then, as we showed before, $\mathbf{v}$ is also an eigenvector corresponding to λ_n, that is, $(\mathbf{K} - \lambda_n \mathbf{M})\mathbf{v} = \mathbf{0}$. Then Lemma 2 with $\lambda = \lambda_n$ demands that

$$\sum_{i,j=1}^{m} a_{ij}(c_i - c_j)^2 = 0. \tag{21}$$

But a_{ij}, given by (16), satisfies $a_{ij} \geqslant 0$ with strict inequality if and only if S_i, S_j are adjacent. Equation (21) implies that if S_i, S_j are adjacent then $c_i = c_j$. This means that if one sign graph, S_i, is omitted in the construction of $\mathbf{v}$ from the sign graphs of $\mathbf{u}$, (that is, $c_i = 0$), then any sign graph S_j adjacent to S_i must also be omitted ($c_j = c_i = 0$). On the other hand if one sign graph S_i is included in $\mathbf{v}$ then any other sign graph S_j adjacent to S_i must be included, and must be included with the same weight as S_i: $c_j = c_i$. This means that in the construction of $\mathbf{v}$ from the sign graphs of $\mathbf{u}$, any connected graph composed of sign graphs of $\mathbf{u}$ must either be included or excluded as a whole. This leads to the following result.

LEMMA 4. *Suppose Conditions* 1 *and* 2 *hold. Suppose* $\mathbf{u}$*, an eigenvector corresponding to* λ_n *has* $n + g$ *sign graphs, where* $g \geqslant 1$*, that is,* $\text{SG}(\mathbf{u}) = n + g$*. These sign graphs may be grouped into* $g + s$ *mutually disjoint connected components* $(C_j)_1^{g+s}$*, and* $s \geqslant 1$.

Proof. If $s < 1$, then there are at most g disjoint connected graphs C_j. If we form a (non-trivial) eigenvector $\mathbf{v}$ from the sign graphs of $\mathbf{u}$, by deleting g of the $n + g$ sign graphs, at least one S_j from each C_j, then $\mathbf{v}$ will include *none* of the S_j; it will be identically zero. This contradiction implies $s \geqslant 1$.

This lemma has a number of implications.

(i) If $\mathbf{u}$ has $m = n + g$ sign graphs, then a connected component C_j can contain at most n sign graphs. For if one contained $n + 1$ sign graphs, then there would be at most $1 + (n + g - n - 1) = g$ connected components. Note that (i) is a somewhat restricted discrete counterpart of Theorem 2.

(ii) If there are n sign graphs in one component C_j, and $n \geqslant 2$, then $g \geqslant 2$. For if n sign graphs are in one component C_j, they must constitute an eigenvector; so too will the remaining $n + g - n = g$ sign graphs. If $n \geqslant 2$, an eigenvector, being orthogonal to $\mathbf{u}_1$, must have at least 2 sign graphs: $g \geqslant 2$.

(iii) If $\mathcal{G}$ is connected and $\mathbf{u}$ has no zeros then, whether λ_n is simple or multiple, $\mathrm{SG}(\mathbf{u}_n) \leqslant n$. For if there are no zero vertices then *all* the sign graphs fall into one component.

To obtain an unrestricted discrete counterpart of Theorem 2 which applies whether λ_n is simple or multiple, we must consider the weak sign graphs rather than the strict sign graphs which we have used. Fiedler (**13**) showed that $\mathbf{u}_n$ can have at most $2(n - 1)$ weak sign graphs. This bound has recently been reduced to n (**14**).

7. Conclusions

Provided that all the triangles are acute-angled, the triangular finite element discretization of the Helmholtz equation exhibits properties that are analogues of those of the continuous equation. The analogues are straightforward for eigenmodes corresponding to simple eigenvalues, but have a restricted form for multiple eigenvalues.

Acknowledgement

We are grateful for the conscientious review by a referee.

References

1. R. Courant and D. Hilbert, *Methods of Mathematical Physics*, Vol. 1 (Interscience, New York 1953).
2. S. Y. Cheng, Eigenfunctions and nodal sets, *Comment. Math. Helvetici* **51** (1976) 43–55.
3. H. Herrmann, Beziehungen zwischen den Eigenwerten und Eigenfunktionen verschiedener Eigenwertprobleme, *Math. Z.* **40** (1935) 221–241.
4. A. Pleijel, Remarks on Courant's nodal line theorem, *Comm. Pure Appl. Math.* **9** (1956) 543–550.
5. D. Jerison and C. Kenig, Unique continuation and absence of positive eigenvalues for Schrödinger operators, *Ann. Math.* **121** (1985) 463–494.
6. F. Pockels, *Über die partielle Differentialgleichung* $\Delta u + k^2 u = 0$ *und deren Auftreten in die mathematischen Physik* (Teubner, Leipzig 1891).
7. H. Zhu, Courant's nodal line theorem and its discrete counterparts, Ph.D. Thesis, University of Waterloo, (2000).
8. R. G. Busacker and T. L. Saaty, *Finite Graphs and Networks: an Introduction with Applications* (McGraw–Hill, New York 1965).
9. M. H. Protter and H. F. Weinberger, *Maximum Principles in Differential Equations* (Springer, New York 1984).
10. A. M. Duval and V. Reiner, Perron–Frobenius type results and discrete versions of nodal domain theorems, *Linear Alg. Appl.* **294** (1999) 259–268.
11. J. Friedman, Some geometric aspects of graphs and their eigenfunctions, *Duke Math. J.* **64** (1993) 487–525.

12. Y. C. de Verdière, Multiplicités des valeurs propres laplaciens discrete et laplaciens continus, *Rend. Mat.* **13** (1993) 433–460.
13. M. Fiedler, A property of eigenvectors of non-negative symmetric matrices and its application to graph theory, *Czech. Math. J.* **25** (1975) 619–633.
14. E. B. Davies, G. M. L. Gladwell, J. Leydold and P. F. Stadler, Discrete nodal domain theorems, *Linear Alg. Appl.* (2001), to appear.