Bounds on decay constants for diffusion through inhomogeneous media

1988 J. Phys. A: Math. Gen. 21 4423

(http://iopscience.iop.org/0305-4470/21/23/023)

J. Phys. A: Math. Gen. 21 (1988) 4423-4441. Printed in the UK

Bounds on decay constants for diffusion through inhomogeneous media

James G Berryman

Courant Institute of Mathematical Sciences, New York University, 251 Mercer Street, New York, NY 10012, USA and Lawrence Livermore National Laboratory†, University of California, PO Box 808 L-156, Livermore, CA 94550, USA

Received 15 April 1988

Abstract. The decay constants for diffusion through inhomogeneous media are known to be proportional to the eigenvalues of the corresponding elliptic operator. A new method of obtaining a hierarchy of upper bounds on sums and products of these eigenvalues as well as the eigenvalues themselves is presented. The first member of this hierarchy is just the usual Rayleigh-Ritz quotient. The other members of the hierarchy are generalised Rayleigh-Ritz quotients which can be derived simply using properties of integrals of the solutions of the diffusion equation. Explicit bounds are presented for the first three eigenvalues, but general methods of obtaining bounds for higher-order eigenvalues are also outlined. For fixed time t, many of the bounds reduce to results given by the classical method of moments. The hierarchy of rigorous variational bounds on the eigenvalues studied may be generated using simple recursion relations based on properties of the characteristic orthogonal polynomials. The conditions on the trial functions used to obtain bounds on eigenvalues higher than the first are much simpler than those required by the traditional Rayleigh-Ritz procedure.

1. Introduction

The diffusion equation to be studied has the form

$$u_t = \nabla(D(\boldsymbol{x})\nabla u). \tag{1}$$

The analysis which follows is valid for arbitrary dimensionality, but we will generally limit discussion to three spatial dimensions. The dependent variable $u(\boldsymbol{x}, t)$ may be interpreted as temperature, density, concentration of chemical species, etc, depending on the particular application. If the initial value

$$u(\boldsymbol{x}, 0) = v(\boldsymbol{x}) \tag{2}$$

inside the finite volume Ω and boundary values

$$u(\boldsymbol{x}, t) = 0 \tag{3}$$

on the boundary $\partial\Omega$, then the solution may be written in terms of the eigenfunctions ψ_n and eigenvalues λ_n of the corresponding elliptic operator

$$\nabla(D(\boldsymbol{x})\nabla\psi_n) = -\lambda_n\psi_n \tag{4}$$

† Permanent address.

as

$$u(\boldsymbol{x}, t) = \sum_{n=1}^{\infty} u_n \psi_n(\boldsymbol{x}) \exp(-\lambda_n t) \tag{5}$$

where the coefficient u_n is given by

$$u_n = \int_\Omega \mathrm{d}x\, v(\boldsymbol{x}) \psi_n(\boldsymbol{x}). \tag{6}$$

The eigenfunctions vanish on the boundary and the orthonormal property

$$\int_\Omega \mathrm{d}x\, \psi_m(\boldsymbol{x}) \psi_n(\boldsymbol{x}) = \delta_{m,n} \tag{7}$$

of the eigenfunctions has been used to obtain (6). We also assume that the initial data are square integrable so that

$$\int_\Omega \mathrm{d}x\, v^2(\boldsymbol{x}) = \sum_{n=1}^{\infty} u_n^2 < \infty. \tag{8}$$

The diffusion coefficient $D(\boldsymbol{x})$ is positive and bounded with enough bounded spatial derivatives to ensure the existence of the various integrals to be defined. The diffusion coefficient is assumed to be inhomogeneous, i.e. spatially varying. However, we will not need to make any of the assumptions commonly made in the theory of composites such as isotropy, randomness, stationarity or statistical homogeneity; similarly, no separation of pertinent length scales will be assumed. Likewise, in contrast to the theory of composites, we specifically exclude piecewise constant diffusion coefficients since these lead to divergent integrals in the hierarchy to be constructed here; such problems require separate treatment.

Estimates of the eigenvalues λ_n are desirable because, as may be seen from (5), they may be used to predict the asymptotic behaviour of the solution of the diffusion equation. The bounds on the eigenvalues presented here are valuable because they may be easily computed during a numerical experiment and will very rapidly converge to good estimates of the eigenvalues during the evolution of the diffusion process, thus eliminating the need for long computations.

Bounds on the eigenvalues may be obtained using the well known Rayleigh–Ritz procedure [1, 2]. First, we define the integrals

$$a(t) = \int \mathrm{d}x\, u^2(\boldsymbol{x}, t) = \sum_n u_n^2 \exp(-2\lambda_n t) \tag{9}$$

$$b(t) = \int \mathrm{d}x\, D(\boldsymbol{x}) |\nabla u|^2 = \sum_n \lambda_n u_n^2 \exp(-2\lambda_n t). \tag{10}$$

Then, the ratio of these two integrals is the well known Rayleigh–Ritz quotient and provides an upper bound on the lowest eigenvalue

$$\lambda_1 \leqslant b/a. \tag{11}$$

(We suppress the time dependence in (11) and throughout much of the paper to simplify the notation since this should cause little confusion. The only time-independent quantities, other than $D(\boldsymbol{x})$ and the initial data, are the true eigenvalues λ_n and the weights u_n.) By considering the integrals

$$c(t)=\int \mathrm{d}x\, u_t^2=\int \mathrm{d}x\,[\nabla(D\nabla u)]^2=\sum \lambda_n^2 u_n^2 \exp(-2\lambda_n t) \tag{12}$$

$$d(t)=\int \mathrm{d}x\, D|\nabla u_t|^2=\int \mathrm{d}x\, D|\nabla[\nabla(D\nabla u)]|^2=\sum \lambda_n^3 u_n^2 \exp(-2\lambda_n t) \tag{13}$$

we will show in § 2 that generalised Rayleigh-Ritz quotients may be obtained to provide bounds on the sum and product

$$\lambda_1+\lambda_* \leqslant \frac{ad-bc}{ac-b^2} \tag{14}$$

$$\lambda_1\lambda_* \leqslant \frac{bd-c^2}{ac-b^2} \tag{15}$$

of the lowest eigenvalue λ_1 and the next lowest eigenvalue λ_* whose eigenfunction has a finite coefficient $u_* \neq 0$. In § 3 we will also find bounds on λ_1 and λ_* that make optimal use of the information contained in the integrals a, b, c, d in the sense that the bounds become exact when there are only two terms in the eigenfunction expansions, and that the bounds are always as tight as any of the other bounds. We show further in § 4 that at fixed time t some of these bounds are the same as the ones given by the classical method of moments, and in § 5 that a hierarchy of rigorous variational bounds on the eigenvalues may be obtained recursively. Some numerical examples are presented in § 6. In § 7 we briefly discuss the previous work on generalised Rayleigh-Ritz quotients and provide several other examples of connections to previous work on numerical methods for estimating eigenvalues and functions of the eigenvalues.

2. Generalised Rayleigh-Ritz quotients

To derive generalised Rayleigh-Ritz quotients, we use a property of solutions of the diffusion equation. This property has to do with the monotonicity of the change in the ratio of various integrals of the solution. To illustrate this property, consider the original Rayleigh-Ritz quotient

$$R_1(t)=b(t)/a(t) \tag{16}$$

as a function of time. The time derivative of the R_1 is then given by

$$\mathrm{d}R_1/\mathrm{d}t=-2(ac-b^2)/a^2 \tag{17}$$

for all $t>0$ and for $t=0$ if the initial data satisfy $u_t(x, 0)=0$. Since the Schwarz inequality implies that (17) is not positive (see equation (A1)) R_1 is a non-increasing function of time. Furthermore, the time derivative vanishes only when $ac=b^2$ which happens if and only if there is exactly one term of the eigenfunction expansion contributing to the integrals a, b, c. Since this special circumstance corresponds to the situation where the initial value $u(x, 0)=v(x)$ is precisely one of the unknown eigenfunctions ψ_n, we may assume this does not occur. Thus, the time derivative $\dot{R}_1$ is strictly negative and the ratio itself is a monotonically decreasing function of time for general initial data.

The Rayleigh-Ritz quotient (16) has the following three properties. (i) It is the positive ratio of combinations of integrals of the solution $u(x, t)$ and its spatial and temporal derivatives. (ii) It is bounded below by a simple but non-trivial function of a finite number of the eigenvalues λ_n. (iii) It is either a constant or a monotonically decreasing function of time. We will define any quotient satisfying these three conditions to be a generalised Rayleigh-Ritz quotient for the diffusion equation. We expect these generalised Rayleigh-Ritz quotients to be useful variational bounds. The variational nature of these quotients is apparent from the fact that they are, in general, monotonically decreasing functions of time for initial data satisfying the boundary conditions.

Next we give three examples of generalised Rayleigh-Ritz quotients according to

$$R_2(t)=\frac{ad-bc}{ac-b^2} \tag{18}$$

$$R_3(t)=\frac{bd-c^2}{ad-bc} \tag{19}$$

$$R_4(t)=\frac{bd-c^2}{ac-b^2}. \tag{20}$$

It follows easily from an examination of the eigenfunction expansions of these ratios (see (A5)-(A7)) that the following inequalities are satisfied for all time:

$$\lambda_1+\lambda_*\leq R_2(t) \tag{21}$$

$$\frac{\lambda_1\lambda_*}{\lambda_1+\lambda_*}\leq R_3(t) \tag{22}$$

$$\lambda_1\lambda_*\leq R_4(t). \tag{23}$$

Furthermore, (A9) and (A12) show that R_2 and R_3 are monotonically decreasing functions of time—unless they are constant for all time. It follows that R_4 is also monotonically decreasing since

$$R_4(t)=R_2(t)R_3(t) \tag{24}$$

and therefore

$$\dot{R}_4(t)=\dot{R}_2(t)R_3(t)+R_2(t)\dot{R}_3(t)\leq 0. \tag{25}$$

Now suppose that for fixed t there exist real constants λ_- and λ_+ satisfying the two equations

$$\lambda_-\lambda_+a-(\lambda_-+\lambda_+)b+c=0 \tag{26}$$

and

$$\lambda_-\lambda_+b-(\lambda_-+\lambda_+)c+d=0. \tag{27}$$

The motivation for considering (26) and (27) will become apparent in the next two sections. That such constants exist follows easily by noting the identities

$$\lambda_\mp=\frac{\lambda_\pm b-c}{\lambda_\pm a-b}=\frac{\lambda_\pm c-d}{\lambda_\pm b-c} \tag{28}$$

which imply that $\lambda_\pm$ are the two solutions of a quadratic equation. Furthermore, taking linear combinations of (26) and (27) shows that

$$\lambda_-+\lambda_+=\frac{ad-bc}{ac-b^2} \tag{29}$$

and

$$\lambda_-\lambda_+ = \frac{bd - c^2}{ac - b^2}. \tag{30}$$

The right-hand sides of (29) and (30) are just R_2 and R_4, while the left-hand sides define two simple curves in the (λ_-, λ_+) plane. The point of intersection of these two curves is the location of the best upper bounds on the first two eigenvalues obtainable from the inequalities (21)-(23). Thus, these inequalities may all be rewritten concisely in terms of the two constants $\lambda_\pm$. A more detailed demonstration that $\lambda_\pm$ are rigorous upper bounds on λ_1 and λ_* will be given in the next section.

3. Optimum bounds on two eigenvalues

To improve on the simple Rayleigh-Ritz bound (11), it is well known that a systematic procedure may be established by adding trial perturbations to the original function (in our case $u(x, 0)$). By making the coefficients of these perturbing functions arbitrary, and then varying them to obtain a minimum value for the ratio, more accurate upper bounds on the lowest eigenvalue may be obtained. This method is systematic and quite general. However, it is clearly advantageous to make careful choices of the trial functions in order to avoid many tedious integrations.

The diffusion equation itself suggests that an optimum choice of trial perturbation is given by u_t. We know that, as $u(x, t)$ evolves from the initial data, the higher terms in the eigenfunction expansion decay exponentially. At some finite time, we expect the solution to be well approximated by the first few terms in the expansion. A perturbation that moves the solution in the right direction is just the first term of the Taylor series expansion:

$$u(x, t) \simeq u(x, 0) + tu_t(x, 0). \tag{31}$$

This observation also provides one motivation for Richardson's iteration method [3-5] for solving the elliptic eigenvalue problem. We will replace t in (31) by a timelike variational coefficient τ and then substitute $u + \tau u_t$ for u in R_1. The resulting Rayleigh-Ritz quotient is given by

$$R_5(\tau) = \frac{b - 2\tau c + \tau^2 d}{a - 2\tau b + \tau^2 c}. \tag{32}$$

The integrals a, b, c, d appearing in (32) are exactly the same as those defined in (9), (10), (12) and (13) but they are evaluated now only at some fixed value of t, say $t = 0$.

The stationary points of (32) occur when the τ variation

$$\frac{dR_5}{d\tau} = -2\frac{(c - \tau d) - (b - \tau c)R_5}{(a - 2\tau b + \tau^2 c)} \tag{33}$$

vanishes. Thus, for values of τ satisfying

$$\frac{c - \tau d}{b - \tau c} = R_5(\tau) \tag{34}$$

(32) achieves its minimum or maximum value.

The quadratic equation that τ satisfies at the stationary points is

$$(ac - b^2) - \tau(ad - bc) + \tau^2(bd - c^2) = 0. \tag{35}$$

The two solutions of the quadratic are given by

$$\tau_{\pm}=\{(ad-bc)\pm[(ad-bc)^2-4(ac-b^2)(bd-c^2)]^{1/2}\}/2(bd-c^2) \qquad (36)$$

which satisfy

$$\tau_-+\tau_+=\frac{ad-bc}{bd-c^2} \qquad (37)$$

and

$$\tau_-\tau_+=\frac{ac-b^2}{bd-c^2}. \qquad (38)$$

It follows from (29), (30), (37) and (38) that

$$\lambda_{\mp}=1/\tau_{\pm}. \qquad (39)$$

Furthermore, it follows from (28), (34) and (39) that

$$R_5(\tau_{\pm})=1/\tau_{\mp}=\lambda_{\pm}. \qquad (40)$$

Thus, the *value* of the function R_5 at one stationary point is given by the reciprocal of the τ *argument* at the other stationary point!

To provide one interpretation of these results, consider the eigenfunction expansions of the numerator and denominator of R_5 and its τ derivative (specifying $t=0$ for simplicity):

$$R_5(\tau)=\frac{\Sigma\, u_m^2\lambda_m(1-\tau\lambda_m)^2}{\Sigma\, u_n^2(1-\tau\lambda_n)^2} \qquad (41)$$

and

$$\frac{\mathrm{d}R_5(\tau)}{\mathrm{d}\tau}=-\frac{\Sigma\, u_k^2u_l^2(\lambda_k-\lambda_l)^2(1-\tau\lambda_k)(1-\tau\lambda_1)}{\Sigma\, u_m^2u_n^2(1-\tau\lambda_m)^2(1-\tau\lambda_n)^2}. \qquad (42)$$

Equation (42) shows that the derivative of R_5 is negative for all $\tau\leqslant 0$. Furthermore, it is also negative for large positive τ. Thus, as τ increases from zero, the graph of R_5 decreases for small τ to a global minimum, then rises to a global maximum, and subsequently falls for larger τ to its asymptotic value $R_5\to d/c$. The global minimum occurs at τ_- and has the value $1/\tau_+=\lambda_-$. Clearly, the global minimum is the upper bound on the lowest eigenvalue λ_1, i.e.

$$\lambda_1\leqslant\lambda_-. \qquad (43)$$

Now consider (41) when $\tau=1/\lambda_1$. We see that

$$\lambda_*\leqslant\frac{\Sigma\, u_m^2\lambda_m(1-\lambda_m/\lambda_1)^2}{\Sigma\, u_n^2(1-\lambda_n/\lambda_1)^2}=R_5(1/\lambda_1) \qquad (44)$$

where λ_* is the next lowest eigenvalue after λ_1 with non-vanishing coefficient u_*. Ordinarily, this eigenvalue will be either λ_3 or λ_2 depending on whether the initial data contain terms only with even parity or not. Inequality (44) shows that at some point the function R_5 attains a value that is an upper bound on λ_*. Note that (44) is an elementary consequence of the well known maximum-minimum property of eigenvalues [6]. The location of this point is unknown unless we know the value of λ_1. So the best bound on λ_* we obtain from this analysis is the value of R_5 at its global maximum, which is

$$\lambda_*\leqslant R_5(\tau_+)=1/\tau_-=\lambda_+. \qquad (45)$$

Thus, we have shown what was postulated at the end of the last section, that λ_- and λ_+ are indeed rigorous upper bounds on the lowest two eigenvalues in the eigenfunction expansion of $u(x, t)$.

To check the optimality of these bounds on the first two eigenvalues, consider the case in which the eigenfunction expansions contain contributions only from $u_1 \neq 0$ and $u_* \neq 0$. Then, the derivative (42) is proportional to

$$\mathrm{d}R_5/\mathrm{d}\tau \propto -u_1^2 u_*^2 (\lambda_1 - \lambda_*)^2 (1 - \tau\lambda_1)(1 - \tau\lambda_*). \tag{46}$$

Clearly, the zeros of (46) occur at precisely $\tau_- = 1/\lambda_*$ and $\tau_+ = 1/\lambda_1$. Furthermore, when (41) is evaluated at these two points, R_5 takes the values λ_1 and λ_*, respectively. Thus, the upper bounds $\lambda_\mp$ found in this section become precise in this special case, and are therefore optimal when the only information being used is the values of the integrals a, b, c, d.

Another interpretation of these results in terms of the classical method of moments is presented in the next section.

Before concluding this section, we want to point out another property of the bounds $\lambda_\pm$. These bounds have been derived here for fixed t. We may now treat the bounds as functions of time to determine how they evolve as the diffusion process progresses. We derive explicit expressions for the time derivatives in the appendix (see (A16)–(A24)) and give an elementary proof that in general

$$\mathrm{d}\lambda_-(t)/\mathrm{d}t \leq 0. \tag{47}$$

In § 5 a more sophisticated analysis shows that $\dot{\lambda}_+ = \mathrm{d}\lambda_+/\mathrm{d}t \leq 0$ also. These observations suggest that both bounds can be improved by including more information about the evolution of the initial data in the variational functional.

4. Method of moments

The results of the last two sections can be understood quite easily in terms of the classical method of moments [7-9]. Recall that the solution to the power moment problem is an approximation to a non-negative measure defined on the line when a finite or infinite number of the moments of the measure are known. Our goal is not the same as that of the power moment problem, but we will see that the same methods apply.

First, note that the Rayleigh–Ritz quotient (16) may be viewed as the solution to the problem

$$\begin{pmatrix} a & b \\ 1 & \lambda \end{pmatrix} \begin{pmatrix} \kappa \\ -1 \end{pmatrix} = 0 \tag{48}$$

where $\kappa = R_1$, i.e. the upper bound on the first eigenvalue λ_1. Furthermore, $\lambda = \kappa$ and therefore the determinant of the matrix in (48) must vanish when (48) is satisfied.

Similarly, (26) and (27) may be rewritten as

$$\begin{pmatrix} a & b & c \\ b & c & d \\ 1 & \lambda & \lambda^2 \end{pmatrix} \begin{pmatrix} \lambda_-\lambda_+ \\ -(\lambda_- + \lambda_+) \\ 1 \end{pmatrix} = 0 \tag{49}$$

where again the determinant of the matrix in (49) must vanish in order for the equation to be satisfied. As was pointed out in the preceding section, λ_- and λ_+ are rigorous upper bounds on λ_1 and λ_2 (or λ_3).

It is well known that (48) and (49) are the first two steps in the solution of the classical moment problem [7]. Therefore, it is natural to consider the succeeding steps in this sequence.

Now introduce the next two integrals in the moment sequence

$$e(t)=\int \mathrm{d}x[\nabla(D\nabla u_t)]^2=\sum \lambda_n^4 u_n^2 \exp(-2\lambda_n t) \tag{50}$$

$$f(t)=\int \mathrm{d}x\, D|\nabla[\nabla(D\nabla u_t)]|^2=\sum \lambda_n^5 u_n^2 \exp(-2\lambda_n t) \tag{51}$$

where u_t is given by (1). Then, the third-order moment problem may be written, in analogy with (48) and (49), as

$$\begin{pmatrix} a & b & c & d \\ b & c & d & e \\ c & d & e & f \\ 1 & \lambda & \lambda^2 & \lambda^3 \end{pmatrix}\begin{pmatrix} \mu_1\mu_2\mu_3 \\ -(\mu_2\mu_3+\mu_3\mu_1+\mu_1\mu_2) \\ \mu_1+\mu_2+\mu_3 \\ -1 \end{pmatrix}=0. \tag{52}$$

We expect to find that μ_1, μ_2, μ_3 are rigorous bounds on $\lambda_1, \lambda_2, \lambda_3$ (or $\lambda_1, \lambda_3, \lambda_5$ if the initial data have purely even parity). To prove that μ_1, μ_2, μ_3 are also rigorous bounds and that succeeding problems in the moment problem sequence produce successively better bounds, we will need to study the properties of the characteristic polynomials.

A sequence of orthogonal polynomials is associated with the preceding problems. These polynomials are defined by the determinants of the matrices in (48), (49), (52) etc. They can also be generated using a simple three-term recursion formula. Let the first two polynomials be defined as

$$P_0(\lambda)=1 \tag{53}$$

and

$$P_1(\lambda)=\det\begin{pmatrix} a & b \\ 1 & \lambda \end{pmatrix}=a\lambda-b. \tag{54}$$

Then we will define the recursion relation

$$D_i\lambda P_i(\lambda)=D_{i-1}P_{i+1}(\lambda)+\alpha_i P_i(\lambda)+\beta_{i-1}P_{i-1}(\lambda) \tag{55}$$

where the leading coefficients D_i are determinants given by

$$D_{-1}=1 \qquad D_0=a \tag{56}$$

$$D_1=\det\begin{pmatrix} a & b \\ b & c \end{pmatrix}=ac-b^2 \tag{57}$$

$$D_2=\det\begin{pmatrix} a & b & c \\ b & c & d \\ c & d & e \end{pmatrix} \tag{58}$$

etc. It is known that the determinants in this sequence are all non-negative [7, 8]. The remaining coefficients α_i and β_i are determined through an orthogonalisation procedure relative to a linear functional $\mathscr{A}$ acting on the space of polynomials in λ with the following characteristics:

$$\mathscr{A}(1)=a \qquad \mathscr{A}(\lambda)=b \qquad \mathscr{A}(\lambda^2)=c \tag{59}$$

etc. For simplicity, we have defined $\mathscr{A}$ operationally, but this functional may also be defined precisely [7] in terms of the weights u_n^2 and the spectrum determined by the eigenvalues λ_n. Then we have

$$\mathscr{A}[\cdot] = \int \mathrm{d}\lambda \sum_{n=1}^{\infty} u_n^2 \delta(\lambda - \lambda_n)[\cdot] \tag{60}$$

and therefore

$$\mathscr{A}[F(\lambda)] = \sum_{n=1}^{\infty} u_n^2 F(\lambda_n). \tag{61}$$

We first check that P_0 and P_1 are orthogonal relative to $\mathscr{A}$ and find

$$\mathscr{A}(P_0 P_1) = a\mathscr{A}(\lambda) - b\mathscr{A}(1) = ab - ba = 0. \tag{62}$$

Then, we require that other polynomials generated by (55) be mutually orthogonal so that

$$\mathscr{A}(P_i P_j) = \mathscr{A}(P_i^2)\delta_{ij}. \tag{63}$$

Multiplying (55) successively by P_{i+1}, P_i, P_{i-1} and then operating with $\mathscr{A}$ produces the equations

$$D_i \mathscr{A}(\lambda P_i P_{i+1}) = D_{i-1} \mathscr{A}(P_{i+1}^2) \tag{64}$$

$$D_i \mathscr{A}(\lambda P_i^2) = \alpha_i \mathscr{A}(P_i^2) \tag{65}$$

and

$$D_i \mathscr{A}(\lambda P_{i-1} P_i) = \beta_{i-1} \mathscr{A}(P_{i-1}^2). \tag{66}$$

From (64)-(66) it follows that

$$\alpha_i = D_i \frac{\mathscr{A}(\lambda P_i^2)}{\mathscr{A}(P_i^2)} \tag{67}$$

and

$$\beta_i = \frac{D_{i-1} D_{i+1}}{D_i} \frac{\mathscr{A}(P_{i+1}^2)}{\mathscr{A}(P_i^2)}. \tag{68}$$

Equation (68) follows from (64) and (66), using (66) to replace the left-hand side of (64). The initial value of the β coefficients is defined to be $\beta_{-1} = 0$. It follows easily from the recursion that the third polynomial in the sequence is

$$P_2(\lambda) = \lambda^2(ac - b^2) - \lambda(ad - bc) + (bd - c^2) = \det \begin{pmatrix} a & b & c \\ b & c & d \\ 1 & \lambda & \lambda^2 \end{pmatrix} \tag{69}$$

as anticipated.

We will now make use of the following well known theorem [7].

Theorem. (i) All the zeros of a real orthogonal polynomial are real and simple. (ii) Any two zeros of the polynomial $P_n(\lambda)$ are separated by a zero of the polynomial $P_{n-1}(\lambda)$ and vice versa.

To prove (i), suppose that the polynomial $P_n(\lambda)$ changes sign only at the points $\mu_1 < \mu_2 < \ldots < \mu_m$, where $m < n$. Then the polynomial

$$Q(\lambda) = P_n(\lambda)(\lambda - \mu_1)(\lambda - \mu_2) \ldots (\lambda - \mu_m)$$

is not identically zero and is non-negative on the real axis, so $\mathscr{A}[Q(\lambda)] > 0$. However, the orthogonality relation requires $\mathscr{A}[Q] = 0$ for $m \leqslant n-2$. It follows that $P_n(\lambda)$ has at least $n-1$ real simple zeros, and therefore that it has n real simple zeros.

To prove (ii), we must first obtain the Christoffel-Darboux formula for the P_n. It is straightforward to show that the formula

$$(\mu - \lambda) \sum_{i=0}^{n-1} \chi_i D_i P_i(\lambda) P_i(\mu) = \gamma_{n-1}[P_{n-1}(\lambda) P_n(\mu) - P_{n-1}(\mu) P_n(\lambda)] \tag{70}$$

is a simple consequence of (55) if the coefficients χ_i and γ_i are defined by

$$\gamma_i = \chi_i D_{i-1} \tag{71}$$

and

$$\chi_i = \frac{1}{D_i \mathscr{A}[P_i^2]}. \tag{72}$$

Note that γ_i and χ_i are both non-negative for all i. Now letting $\mu \to \lambda$ in (70), we find

$$\sum_{i=0}^{n-1} \chi_i D_i P_i^2(\lambda) = \gamma_{n-1}[P_{n-1}(\lambda) P_n'(\lambda) - P_{n-1}'(\lambda) P_n(\lambda)] > 0. \tag{73}$$

If μ_i and μ_{i+1} are neighbouring zeros of $P_n(\lambda)$, then the values of the derivatives $P_n'(\mu_i)$ and $P_n'(\mu_{i+1})$ must have opposite sign. Equation (73) shows on the other hand that

$$P_{n-1}(\mu_i) P_n'(\mu_i) > 0 \tag{74}$$

and

$$P_{n-1}(\mu_{i+1}) P_n'(\mu_{i+1}) > 0 \tag{75}$$

so $P_{n-1}(\mu_i)$ and $P_{n-1}(\mu_{i+1})$ must also be of opposite sign. A similar result obtains when zeros of $P_{n-1}(\lambda)$ are considered. The interlacing property of the zeros of the polynomials quoted in the theorem then follows directly from these facts.

Thus the theorem shows that, if $\mu_1^{(n)}, \mu_2^{(n)}, \ldots,$ are the zeros of $P_n(\lambda)$ and $\mu_1^{(n-1)}, \mu_2^{(n-1)}, \ldots,$ are the zeros of $P_{n-1}(\lambda)$, then

$$\mu_1^{(n)} \leqslant \mu_1^{(n-1)} \leqslant \mu_2^{(n)} \leqslant \mu_2^{(n-1)} \leqslant \mu_3^{(n)} \leqslant \ldots. \tag{76}$$

These relationships hold at any fixed value of the time parameter t. Furthermore, it follows easily from induction on (76) that

$$\lambda_1 \leqslant \mu_1^{(n)} \leqslant \mu_1^{(n-1)} \leqslant \ldots \leqslant \mu_1^{(1)} \tag{77}$$

$$\lambda_2 \leqslant \lambda_* \leqslant \mu_2^{(n)} \leqslant \mu_2^{(n-1)} \leqslant \ldots \leqslant \mu_2^{(2)} \tag{78}$$

$$\lambda_3 \leqslant \lambda_{**} \leqslant \mu_3^{(n)} \leqslant \mu_3^{(n-1)} \leqslant \ldots \leqslant \mu_3^{(3)} \tag{79}$$

etc. The next lowest eigenvalues λ_*, λ_{**} may coincide with λ_2, λ_3 or with λ_3, λ_5 or with some other pair of eigenvalues depending on the initial data. Thus, the zeros from the sequence of orthogonal polynomials form a convergent sequence of rigorous bounds on some subset of the eigenvalues.

Equations (76)–(79) show the relations existing among the various zeros of the characteristic polynomials and the eigenvalues at a fixed value of time t. In the next section we show how these zeros change as a function of time.

5. Hierarchy of variational bounds

Now we wish to study the higher-order generalised Rayleigh–Ritz quotients. From (52), we easily find expressions for the sums and products of the next three bounds in this sequence:

$$\mu_1\mu_2\mu_3 = D_2^{-1} \det \begin{pmatrix} b & c & d \\ c & d & e \\ d & e & f \end{pmatrix} \tag{80}$$

$$\mu_2\mu_3 + \mu_3\mu_1 + \mu_1\mu_2 = D_2^{-1} \det \begin{pmatrix} a & c & d \\ b & d & e \\ c & e & f \end{pmatrix} \tag{81}$$

$$\mu_1 + \mu_2 + \mu_3 = D_2^{-1} \det \begin{pmatrix} a & b & d \\ b & c & e \\ c & d & f \end{pmatrix}. \tag{82}$$

That these quotients are positive and bounded below by the corresponding functions of the eigenvalues $\lambda_1, \lambda_2, \lambda_3$ is easy to check. Our definition of a generalised Rayleigh–Ritz quotient requires in addition that the ratios be decreasing functions of time. We will now show not only that the expressions (80)–(82) and their higher-order generalisations (i.e. the elementary symmetric functions on the zeros of the polynomials) are decreasing functions of time but also the stronger result that each of the factors μ_1, μ_2, μ_3 and their generalisations are also decreasing functions of time.

We now introduce the kernel polynomial of order n:

$$K_n(\lambda, \mu) = \sum_{i=0}^{n} \frac{P_i(\lambda)P_i(\mu)}{\mathscr{A}[P_i^2]}. \tag{83}$$

One identity satisfied by the kernel polynomial was presented in (70). It is straightforward to check that the kernel polynomial may also be represented explicitly as the ratio of two determinants

$$K_n(\lambda, \mu) = -D_n^{-1} \det \begin{pmatrix} a & b & c & \dots & \mathscr{A}(\lambda^n) & 1 \\ b & c & d & \dots & \mathscr{A}(\lambda^{n+1}) & \mu \\ c & d & e & \dots & \mathscr{A}(\lambda^{n+2}) & \mu^2 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ \mathscr{A}(\lambda^n) & \mathscr{A}(\lambda^{n+1}) & \mathscr{A}(\lambda^{n+2}) & \dots & \mathscr{A}(\lambda^{2n}) & \mu^n \\ 1 & \lambda & \lambda^2 & \dots & \lambda^n & 0 \end{pmatrix} \tag{84}$$

since the kernel polynomial is completely determined by the fact that for any polynomial $\tilde{P}(\lambda)$ of degree $\leqslant n$

$$\mathscr{A}[\tilde{P}(\lambda)K_n(\lambda, \mu)] = \tilde{P}(\mu). \tag{85}$$

For example,

$$\mathscr{A}[\lambda^m K_n(\lambda,\mu)] = -D_n^{-1}\det\begin{pmatrix} a & b & c & \dots & \mathscr{A}(\lambda^n) & 1 \\ b & c & d & \dots & \mathscr{A}(\lambda^{n+1}) & \mu \\ c & d & e & \dots & \mathscr{A}(\lambda^{n+2}) & \mu^2 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ \mathscr{A}(\lambda^n) & \mathscr{A}(\lambda^{n+1}) & \mathscr{A}(\lambda^{n+2}) & \dots & \mathscr{A}(\lambda^{2n}) & \mu^n \\ \mathscr{A}(\lambda^m) & \mathscr{A}(\lambda^{m+1}) & \mathscr{A}(\lambda^{m+2}) & \dots & \mathscr{A}(\lambda^{m+n}) & 0 \end{pmatrix} = \mu^m. \tag{86}$$

The following identity will play an important role in the analysis of the time dependence of the variational bounds:

$$\frac{d^{n-1}K_n(\lambda,0)}{d\mu^{n-1}} = \sum_{i=n-1}^{n} \frac{P_i(\lambda)[d^{n-1}P_i(0)/d\mu^{n-1}]}{\mathscr{A}[P_i^2]}. \tag{87}$$

Equation (87) follows from (83) using the fact that for $i \leqslant n-2$ the μ derivatives of order i or higher vanish identically.

Now we consider the time dependence of the zeros of the characteristic polynomials. In general, we have

$$P_n(\mu_i^{(n)}) = 0. \tag{88}$$

The polynomial is an implicit function of the moments a, b, c, etc. Taking the total time derivative of (88), we have

$$\dot{\mu}_i^{(n)} P_n'(\mu_i^{(n)}) = 2\det\begin{pmatrix} a & b & c & \dots & \mathscr{A}(\lambda^n) \\ b & c & d & \dots & \mathscr{A}(\lambda^{n+1}) \\ c & d & e & \dots & \mathscr{A}(\lambda^{n+2}) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \mathscr{A}(\lambda^{n-2}) & \mathscr{A}(\lambda^{n-1}) & \mathscr{A}(\lambda^n) & \dots & \mathscr{A}(\lambda^{2n-2}) \\ \mathscr{A}(\lambda^n) & \mathscr{A}(\lambda^{n+1}) & \mathscr{A}(\lambda^{n+2}) & \dots & \mathscr{A}(\lambda^{2n}) \\ 1 & \mu_i^{(n)} & [\mu_i^{(n)}]^2 & \dots & [\mu_i^{(n)}]^n \end{pmatrix}. \tag{89}$$

We have used the fact that $\dot{\mathscr{A}}(\lambda^{n-1}) = -2\mathscr{A}(\lambda^n)$ for all $t > 0$ to simplify the determinants. (If the present method is to be used to obtain analytical estimates, some care must be taken to ensure that $\dot{\mathscr{A}}(\lambda^{n-1}) = -2\mathscr{A}(\lambda^n)$ is also true at $t=0$. This condition places admissibility constraints on time derivatives of the initial data at $t=0$. If we are bounding n eigenvalues, we have $n-1$ conditions on the time derivatives of u at $t=0$.)

Now from (84) it follows that

$$\frac{d^{n-1}K_n(\lambda,0)}{d\mu^{n-1}} = -(n-1)!\,D_n^{-1} \times\det\begin{pmatrix} a & b & c & \dots & \mathscr{A}(\lambda^n) \\ b & c & d & \dots & \mathscr{A}(\lambda^{n+1}) \\ c & d & e & \dots & \mathscr{A}(\lambda^{n+2}) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \mathscr{A}(\lambda^{n-2}) & \mathscr{A}(\lambda^{n-1}) & \mathscr{A}(\lambda^n) & \dots & \mathscr{A}(\lambda^{2n-2}) \\ \mathscr{A}(\lambda^n) & \mathscr{A}(\lambda^{n+1}) & \mathscr{A}(\lambda^{n+2}) & \dots & \mathscr{A}(\lambda^{2n}) \\ 1 & \lambda & \lambda^2 & \dots & \lambda^n \end{pmatrix}. \tag{90}$$

Equation (86) shows in addition that

$$\frac{\mathrm{d}^{n-1}K_n(\mu_i^{(n)},0)}{\mathrm{d}\mu^{n-1}}=\frac{P_{n-1}(\mu_i^{(n)})[\mathrm{d}^{n-1}P_{n-1}(0)/\mathrm{d}\mu^{n-1}]}{\mathscr{A}[P_{n-1}^2]}. \tag{91}$$

Equations (90) and (91) together show that (89) may be rewritten as

$$\dot{\mu}_i^{(n)}=-\frac{2D_{n-2}D_nP_{n-1}(\mu_i^{(n)})}{\mathscr{A}[P_{n-1}^2]P_n'(\mu_i^{(n)})} \tag{92}$$

using the fact that

$$\frac{\mathrm{d}^{n-1}P_{n-1}(0)}{\mathrm{d}\mu^{n-1}}=(n-1)!\,D_{n-2}. \tag{93}$$

Equation (92) is a general expression for the time derivative of the ith zero of the nth characteristic polynomial. The sign of $\dot{\mu}_i^{(n)}$ depends only on the sign of the ratio P_{n-1}/P_n'.

We may obtain a general result concerning the sign of the first derivative of the P_n by induction on (74) and (75). In particular, we know that $P_0=1$, $P_0'=0$ and $P_1'>0$ everywhere. From this and the previous results, it follows that $P_1'(\mu_1^{(1)})>0$, $P_2'(\mu_1^{(2)})<0$, $P_3'(\mu_1^{(3)})>0$, etc. Thus, in general, we have

$$(-1)^{n+1}P_n'(\mu_1^{(n)})>0. \tag{94}$$

Using the interlacing property of the zeros and the alternating signs of the derivatives then gives the simple result that

$$(-1)^{n+i}P_n'(\mu_i^{(n)})>0 \tag{95}$$

while (74) and (94) then show that

$$(-1)^{n+i}P_{n-1}(\mu_i^{(n)})>0. \tag{96}$$

Thus, the ratio appearing in (92) satisfies

$$P_{n-1}(\mu_i^{(n)})/P_n'(\mu_i^{(n)})>0. \tag{97}$$

Alternatively, we could have simply noted that, if the product (74) is positive, then the ratio is also positive.

Thus, when we consider the zeros $\mu_i^{(n)}$ to be functions of time, we find that

$$\dot{\mu}_i^{(n)}(t)\leqslant 0 \tag{98}$$

for all i and all n. We have therefore proven that expressions such as (80)-(82) and generalisations are indeed generalised Rayleigh-Ritz quotients as we had supposed. So, in addition to the hierarchy of bounds on eigenvalues given by (77)-(79) at fixed time, we also have another hierarchy of bounds for fixed polynomial degree n as the time parameter marches forward, say in units of Δt for a numerical simulation. The resulting equations are

$$\lambda_1\leqslant\mu_1^{(n)}(m\Delta t)\leqslant\mu_1^{(n)}((m-1)\Delta t)\leqslant\ldots\leqslant\mu_1^{(n)}(0) \tag{99}$$

$$\lambda_2\leqslant\lambda_*\leqslant\mu_2^{(n)}(m\Delta t)\leqslant\mu_2^{(n)}((m-1)\Delta t)\leqslant\ldots\leqslant\mu_2^{(n)}(0) \tag{100}$$

$$\lambda_3\leqslant\lambda_{**}\leqslant\mu_3^{(n)}(m\Delta t)\leqslant\mu_3^{(n)}((m-1)\Delta t)\leqslant\ldots\leqslant\mu_3^{(n)}(0) \tag{101}$$

providing a second hierarchy of variational bounds on the eigenvalues.

6. Applications

6.1. *Extrapolation*

The emphasis of the preceding analysis has been directed toward obtaining rigorous bounds on the eigenvalues. We will now relax the rigour of our objectives and attempt to find direct estimates of the eigenvalues. Although these estimates may fail to be useful in some pathological cases, we may hope such cases can be eliminated by a careful choice of initial data.

Consider the approximation

$$\mu_i^{(n)}(t) \simeq \mu_i^{(n)}(0) + [\lambda_i - \mu_i^{(n)}(0)][1 - \exp(-\varepsilon_i^{(n)} t)]. \tag{102}$$

Equation (102) is based on our knowledge that the eigenvalue bounds are approaching the eigenvalues asymptotically at an exponential rate. If we knew the value of the (approximately constant) exponent $\varepsilon_i^{(n)}$ as well as $\mu_i^{(n)}$, then we could take the time derivative of (102):

$$\dot{\mu}_i^{(n)}(t) \simeq \varepsilon_i^{(n)}[\lambda_i - \mu_i^{(n)}(0)] \exp(-\varepsilon_i^{(n)} t) \tag{103}$$

and solve for λ_i at $t = 0$:

$$\lambda_i \simeq \mu_i^{(n)}(0) + \dot{\mu}_i^{(n)}(0)/\varepsilon_i^{(n)}. \tag{104}$$

We can estimate the exponent itself by taking a second time derivative of $\mu_i^{(n)}$; (104) then becomes

$$\lambda_i \simeq \mu_i^{(n)}(0) - [\dot{\mu}_i^{(n)}(0)]^2/\ddot{\mu}_i^{(n)}(0). \tag{105}$$

As an example, consider

$$\mu_1^{(1)} = \kappa = b/a \tag{106}$$

$$\dot{\mu}_1^{(1)} = -2(ac - b^2)/a^2 \tag{107}$$

$$\ddot{\mu}_1^{(1)} = 4[a(ad - bc) - 2b(ac - b^2)]/a^3. \tag{108}$$

Then the extrapolation formula (105) becomes (using the notation from (48) and (49)):

$$\lambda_1 \simeq \tilde{\lambda}_1 = \frac{\lambda_- \lambda_+ - \kappa^2}{\lambda_- + \lambda_+ - 2\kappa}. \tag{109}$$

As $t \to \infty$ we see from (109) that

$$\tilde{\lambda}_1 \to \frac{\lambda_1 \lambda_* - \lambda_1^2}{\lambda_1 + \lambda_* - 2\lambda_1} \to \lambda_1. \tag{110}$$

Equation (109) is itself neither a bound nor a generalised Rayleigh–Ritz quotient because, when we compute the time derivative of (109), we find

$$\mathrm{d}\tilde{\lambda}_1/\mathrm{d}t = [2\dot{\kappa}(\kappa - \lambda_-)(\kappa - \lambda_+) + \dot{\lambda}_-(\lambda_+ - \kappa)^2 + \dot{\lambda}_+(\kappa - \lambda_-)^2]/(\lambda_- + \lambda_+ - \kappa)^2. \tag{111}$$

The sign of (111) is not generally known unless $(\kappa - \lambda_+) \geq 0$.

Extrapolation formulae such as (109) are well known [10, 11]. Shanks [12] presents a systematic method for producing more sophisticated formulae of this type, but we will not pursue the generalisations of (109) here.

6.2. Examples

To provide a one-dimensional example, consider

$$D(x) = 1 + \alpha \cos^2 \pi x \tag{112}$$

with initial data

$$v(x) = \sin \pi x. \tag{113}$$

Then, a straightforward calculation shows that

$$a(0) = \tfrac{1}{2} \tag{114}$$

$$b(0) = \pi^2(\tfrac{1}{2} + \tfrac{3}{8}\alpha) \tag{115}$$

$$c(0) = \pi^4(\tfrac{1}{2} + \tfrac{3}{4}\alpha + \tfrac{9}{16}\alpha^2) \tag{116}$$

$$d(0) = \pi^6[(\tfrac{1}{2} + \tfrac{3}{8}\alpha)(1 - 6\alpha)^2 + \tfrac{9}{4}\alpha(3 + \tfrac{5}{2}\alpha)(1 - 6\alpha) + \tfrac{405}{16}\alpha^2(1 + \tfrac{7}{8}\alpha)]. \tag{117}$$

With $\alpha = \frac{1}{6}$, we find $a = 0.5$, $b = 5.55$, $c = 62.40$, $d = 774.6$ and the bounds become $\kappa = 11.10$, $\lambda_- = 11.09$ for λ_1 and $\lambda_+ = 96.24$ for λ_3. The extrapolated value $\tilde{\lambda}_1$ differs negligibly from λ_-, since λ_+ is so large. For comparison, note that since $1 \leq D(x) \leq \frac{7}{6}$ the maximum principle [13] shows that $\pi^2 = 9.87 \leq \lambda_1 \leq \frac{7}{6}\pi^2 = 11.51$. Using the sine functions $\sin n\pi x$ as an expansion set, Gerschgorin's theorem [5] shows also that $9\pi^2 = 88.83 \leq \lambda_3 \leq \frac{63}{6}\pi^2 = 103.63$. Thus, we have obtained an apparently quite accurate bound on λ_3 without needing to find an approximation to the corresponding eigenfunction. Using a Galerkin method [14] with basis functions $\sin \pi x$ and $\sin 3\pi x$, we obtain results for the eigenvalues virtually identical to those obtained from the generalised Rayleigh-Ritz method using just $\sin \pi x$ as the trial function (initial data). Finally, note that this example was carefully chosen so that

$$u_t = \nabla(D(x)\nabla u) = 0 \tag{118}$$

on the boundaries; this condition is required as it is used in the proofs of the monotonic convergence of the polynomial zeros to the eigenvalues. Equation (118) is the only additional admissibility condition required by the method when trying to bound two eigenvalues.

Other applications of the method of moments to diffusion problems have been presented by Vorobyev [9].

7. Discussion and conclusions

Generalised Rayleigh-Ritz quotients for diffusion problems were first introduced by Berryman and Holland [15] in studies of the asymptotic behaviour of non-linear diffusion when the diffusion coefficient is a power $(\delta - 1)$ of the dependent variable, i.e. for the porous media equations:

$$u_t = \nabla(D(u)\nabla u).$$

The diffusion coefficient is then given by

$$D(u) = \delta u^{\delta - 1}$$

and the appropriate generalised Rayleigh-Ritz quotient is

$$R(t) = \frac{\int dx |\nabla u^\delta|^2}{(\int dx\, u^{\delta+1})^p}$$

where the exponent $p = 2\delta/(\delta+1)$. The time derivative of $R(t)$ is therefore non-positive since

$$\left(\int \mathrm{d}x\, u^{\delta}\nabla^2 u^{\delta}\right)^2 \leq \int \mathrm{d}x\, u^{\delta+1} \int \mathrm{d}x\, u^{\delta-1}|\nabla^2 u^{\delta}|^2$$

follows easily from the Schwarz inequality. The generalised Rayleigh–Ritz quotient is useful for studies of the asymptotic behaviour of the solutions of the porous media equations for all δ in the range $0<\delta<\infty$. Note that the special value $\delta = 1$ corresponds to linear diffusion as considered here in §§1–6.

The results presented here are valid for arbitrary dimensionality and arbitrary boundary shape. In fact, the boundary need not be at all simple—the results apply in particular to the problem of reaction-controlled diffusion where there are many absorbing surfaces scattered among the diffusing species. The method is also not restricted to Dirichlet boundary conditions. If (3) is replaced by a no-flux (Neumann) boundary condition $\hat{n}\cdot\nabla u = 0$, then the integral $\int_{\Omega} \mathrm{d}x\, u(x, t)$ is constant and

$$u(x, t) \to u_0 = \Omega^{-1} \int_{\Omega} \mathrm{d}x\, u(x, 0).$$

Then it is not difficult to show that the ratio

$$R(t) = \frac{\int \mathrm{d}x\, D|\nabla u|^2}{\int \mathrm{d}x (u-u_0)^2}$$

is a generalised Rayleigh–Ritz coefficient for this problem.

The existence of two distinct hierarchies of bounds suggests that there may exist some simple relationship between the two. We will not explore this question in depth here, but simply note that the analysis of § 3 has already shown that $\mu_1^{(2)}(0) \simeq \mu_1^{(1)}(\tau_-)$. It may very well be possible to establish a detailed correspondence between the two sets of bounds.

When a large number of eigenvalues is being sought, numerical methods for implementing the ideas described here will be required. For such circumstances, it may be useful to point out that the escalator method [16, 17] for finding eigenvalues of systems of equations is closely related to the methods developed here. In the escalator method, the zeros of the polynomial P_{n-1} are used to accelerate the accurate computation of the zeros of P_n.

It was pointed out in § 3 that the bound on the second eigenvalue obtained from (44) is an elementary consequence of the maximum-minimum property of eigenvalues [6]. Indeed, the existence of a bound on the nth eigenvalue, using only the information contained in the characteristic polynomial $P_n(\lambda)$, is also guaranteed by the same property.

The first variational characterisations of elementary symmetric functions on the eigenvalues of positive-definite matrices are apparently due to Fan [18]. Beckenbach and Bellman [19] summarise these results and also present a list of papers containing generalisations of these ideas. The elementary symmetric functions also play a role in the numerical methods for estimating eigenvalues of finite systems of equations due to Aitken [10].

Although the result (98) appears to be new, the method of the derivation has much in common with a method used by Markov (see Gantmacher [20]) to prove a theorem on domains of stability.

We conclude that a new method of obtaining bounds on the eigenvalues of an elliptic operator has been developed. The method has much in common with the classical method of moments, but the relationship discovered between the zeros of the characteristic polynomials and the eigenvalues as the diffusion process evolves appears to be new. The generalised Rayleigh-Ritz quotients presented here have the advantage that information about eigenvalues can be obtained without constructing approximations to the corresponding eigenfunctions. In contrast, the traditional method of applying the Rayleigh-Ritz method to find eigenvalues of order higher than the first requires a careful orthogonalisation of the trial eigenfunction with respect to the lower-order eigenfunctions constructed by the method. The only penalty that must be paid to use the generalised Rayleigh-Ritz method is that the trial fields (or initial data) must satisfy as many of the conditions $\dot{a}(0) = -2b(0)$, $\dot{b}(0) = -2c(0)$, etc, as are actually needed at the order to which we are working. These conditions merely require that the trial functions satisfy the conditions $u_t = 0$, $u_{tt} = 0$, etc, on the boundary at $t = 0$. Such trial functions can always be constructed.

Acknowledgments

I thank M Avellaneda, G W Milton and G C Papanicolaou for very helpful conversations. This work was supported in part by the Air Force Office of Scientific Research through grants AFOSR 85-0017 and AFOSR 86-0352, and also in part under the auspices of the US Department of Energy by the Lawrence Livermore National Laboratory under contract no W-7405-ENG-48.

Appendix. Some relevant identities and inequalities

From the Schwarz inequality it is straightforward to show that the integrals defined in (9), (10), (12) and (13) satisfy the following inequalities:

$$ac - b^2 \geqslant 0 \tag{A1}$$

$$bd - c^2 \geqslant 0. \tag{A2}$$

Using (A1) and (A2) it is clear that

$$abcd \geqslant b^2c^2 \tag{A3}$$

from which it follows easily that

$$ad - bc \geqslant 0. \tag{A4}$$

These combinations of integrals occur repeatedly in the analysis of the text, so it is worthwhile to point out the relationship between these factors and the eigenfunction expansions. We find easily that

$$ac - b^2 = \tfrac{1}{2}\sum u_m^2 u_n^2 (\lambda_m - \lambda_n)^2 \exp[-2(\lambda_m + \lambda_n)t] \tag{A5}$$

$$bd - c^2 = \tfrac{1}{2}\sum u_m^2 u_n^2 (\lambda_m - \lambda_n)^2 \lambda_m \lambda_n \exp[-2(\lambda_m + \lambda_n)t] \tag{A6}$$

$$ad - bc = \tfrac{1}{2}\sum u_m^2 u_n^2 (\lambda_m - \lambda_n)^2 (\lambda_m + \lambda_n) \exp[-2(\lambda_m + \lambda_n)t]. \tag{A7}$$

Now define the quotient

$$R_2(t) = \frac{ad - bc}{ac - b^2}. \tag{A8}$$

Then, we find at $t=0$ that

$$\frac{\mathrm{d}R_2}{\mathrm{d}t}=-\tfrac{1}{4}\sum u_k^2u_l^2u_m^2u_n^2(\lambda_k-\lambda_l)^2(\lambda_m-\lambda_n)^2(\lambda_k+\lambda_l-\lambda_m-\lambda_n)^2/(ac-b^2)^2\leqslant 0 \tag{A9}$$

showing that $R_2(t)$ is a monotonically decreasing function of time unless there are only two non-vanishing coefficients u_1 and u_*, in which case the ratio takes the constant value

$$R_2(t)=\lambda_1+\lambda_*. \tag{A10}$$

Similarly, defining the quotient

$$R_3(t)=\frac{bd-c^2}{ad-bc} \tag{A11}$$

we find that

$$\frac{\mathrm{d}R_3}{\mathrm{d}t}=-\frac{c}{3}\sum u_l^2u_m^2u_n^2(\lambda_n-\lambda_l)^2(\lambda_l-\lambda_m)^2(\lambda_m-\lambda_n)^2/(ad-bc)^2\leqslant 0. \tag{A12}$$

Equality applies in (A12) if and only if the ratio takes the constant value

$$R_3=\frac{\lambda_1\lambda_*}{\lambda_1+\lambda_*}. \tag{A13}$$

It follows from (A8) and $\dot{R}_2\leqslant 0$ that

$$\frac{ae-c^2}{ad-bc}\geqslant\frac{ad-bc}{ac-b^2} \tag{A14}$$

where e is defined in (50). Similarly, from (A11) and $\dot{R}_3\leqslant 0$, we obtain

$$\frac{be-cd}{ad-bc}\geqslant\frac{bd-c^2}{ac-b^2}. \tag{A15}$$

Now we will compute the time derivatives of the bounds $\lambda_\pm$. From (49), we have

$$\det\begin{pmatrix} a & b & c\\ b & c & d\\ 1 & \lambda & \lambda^2\end{pmatrix}=0. \tag{A16}$$

Taking the time derivative of (A16), we have

$$\dot{\lambda}=2\det\begin{pmatrix} a & b & c\\ c & d & e\\ 1 & \lambda & \lambda^2\end{pmatrix}\left[\det\begin{pmatrix} a & b & c\\ b & c & d\\ 0 & 1 & 2\lambda\end{pmatrix}\right]^{-1} \tag{A17}$$

where we have used the fact that $\dot{a}=-2b$, $\dot{b}=-2c$, $\dot{c}=-2d$, etc (which is certainly true for all $t>0$ and can be guaranteed also for $t=0$ by a suitable choice of initial data) to simplify the determinants. Equation (A17) may be rewritten as

$$\dot{\lambda}=2\frac{\lambda^2(ad-bc)-\lambda(ae-c^2)+(bd-c^2)}{2\lambda(ac-b^2)-(ad-bc)} \tag{A18}$$

It follows from (36) and (39) that

$$\lambda_\pm-\frac{(ad-bc)}{2(ac-b^2)}=\pm\frac{1}{2(ac-b^2)}[(ad-bc)^2-4(ac-b^2)(bd-c^2)]^{1/2}. \tag{A19}$$

Thus, the denominator of (A18) is positive for λ_+ and negative for λ_-. The polynomial in the numerator is proportional to

$$\tilde{P}(\lambda)=\lambda^2-\lambda(ae-c^2)/(ad-bc)+(be-cd)/(ad-bc) \tag{A20}$$

and has its minimum value at $\tilde{\lambda}_{\min}=(ae-c^2)/2(ad-bc)$, while the polynomial

$$P(\lambda)=\lambda^2-\lambda(ad-bc)/(ac-b^2)+(bd-c^2)/(ac-b^2) \tag{A21}$$

that determines $\lambda_\pm$ has its minimum at $\lambda_{\min}=(ad-bc)/2(ac-b^2)$. These two polynomials cross at most once and the relative location of the minima is given by

$$\lambda_{\min}\leqslant\tilde{\lambda}_{\min} \tag{A22}$$

which follows from (A14). Therefore, it is clear that

$$\tilde{P}(\lambda_-)\geqslant 0. \tag{A23}$$

Inequalities (A19) and (A23) together show that

$$\mathrm{d}\lambda_-(t)/\mathrm{d}t\leqslant 0. \tag{A24}$$

A more sophisticated method of analysis (see § 5) shows that $\tilde{P}(\lambda_+)\leqslant 0$ and therefore that $\dot{\lambda}_+\leqslant 0$ also.

References

[1] Collatz L 1960 *The Numerical Treatment of Differential Equations* (Berlin: Springer) pp 227-49

[2] Weinberger H F 1956 *Comm. Pure Appl. Math.* **9** 613; 1958 *Pac. J. Math.* **8** 339; 1962 *Variational Methods for Eigenvalue Problems* University of Minnesota Lecture Notes pp 110-27

[3] Frankel S P 1950 *Math. Tables Aids Comp.* **4** 65

[4] Birkhoff G and Lynch R E 1984 *Numerical Solution of Elliptic Problems* (Philadelphia, PA: SIAM) pp 144-8

[5] Varga R S 1962 *Matrix Iterative Analysis* (Englewood Cliffs, NJ: Prentice-Hall) pp 16-7, 141, 159

[6] Courant R and Hilbert D 1953 *Methods of Mathematical Physics* vol I (New York: Interscience) pp 132-4

[7] Akhiezer N I 1965 *The Classical Moment Problem* (New York: Hafner) pp 1-15, 60-7

[8] Lonke A 1971 *J. Math. Phys.* **12** 2422

[9] Vorobyev Yu V 1965 *Method of Moments in Applied Mathematics* (New York: Gordon and Breach) pp 44-7, 114-5

[10] Aitken A C 1926 *Proc. R. Soc. Edin.* **46** 289

[11] Greenberg M D 1978 *Foundations of Applied Mathematics* (Englewood Cliffs, NJ: Prentice-Hall) pp 38-41

[12] Shanks D 1955 *J. Math. and Phys.* **34** 1

[13] Protter M H and Weinberger H F 1984 *Maximum Principles in Differential Equations* (Berlin: Springer) pp 1-194

[14] Kantorovich L V and Krylov V I 1958 *Approximate Methods of Higher Analysis* (New York: Interscience) pp 258-62

[15] Berryman J G and Holland C J 1978 *J. Math. Phys.* **19** 2476

[16] Morris J 1947 *The Escalator Method in Engineering Vibration Problems* (New York: Wiley) pp 111-30

[17] Milne W E 1953 *Numerical Solution of Differential Equations* (New York: Wiley) pp 156-69, 189-90

[18] Fan K 1951 *Am. Math. Mon.* **58** 194

[19] Beckenbach E F and Bellman R 1961 *Inequalities* (Berlin: Springer) pp 76-9

[20] Gantmacher F R 1960 *The Theory of Matrices* vol II (New York: Chelsea) pp 240-8