Available online at www.sciencedirect.com

SCIENCE DIRECT®

Journal of Computational Physics 206 (2005) 302–310

JOURNAL OF COMPUTATIONAL PHYSICS

www.elsevier.com/locate/jcp

A Chebyshev/rational Chebyshev spectral method for the Helmholtz equation in a sector on the surface of a sphere: defeating corner singularities

John P. Boyd *

Department of Atmospheric, Oceanic and Space Science, University of Michigan, 2455 Hayward Avenue, Ann Arbor, MI 48109-2143, USA

Received 1 June 2004; received in revised form 12 October 2004; accepted 10 December 2004
Available online 22 January 2005

Abstract

When the boundaries of a domain meet at an angle, the solutions to an elliptic partial differential equation will usually be singular at the corner. Using the example of the Helmholtz equation on the surface of a sphere in a domain bounded by meridians, we show how corner singularities can be defeated by mapping the corner to infinity. By applying a Chebyshev series in longitude and a rational Chebyshev series in the "Mercator" coordinate, $y = \operatorname{arctanh}(\cos(\text{colatitude}))$, we obtain an exponential rate of convergence despite the corner singularities.

Keywords: Pseudospectral; Chebyshev polynomial; Rational Chebyshev functions; Corner singularities

1. Introduction

In spherical coordinates (λ, μ), where λ is longitude and μ is the cosine of colatitude, the Helmholtz equation on the surface of a sphere is

$$\frac{1}{1-\mu^2}u_{\lambda\lambda} + (1-\mu^2)u_{\mu\mu} - 2\mu u_\mu + k^2 u = f(\lambda, \mu), \tag{1}$$

where k is a constant. The domain is the sector bounded by the meridians $\lambda = 0$, Ξ, where Ξ is a constant as shown in Fig. 1. Similar sectorial wave problems arise in ocean tides [7,10,19,21] , but for expository

* Tel.: +1 734 764 3338; fax: +1 734 764 5137.
E-mail address: jpboyd@umich.edu.

0021-9991/$ - see front matter
doi:10.1016/j.jcp.2004.12.006

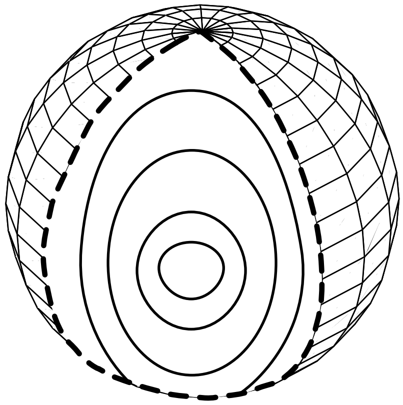

Fig. 1. Schematic of a sectorial domain on the surface of a sphere. The thick dashed lines are the boundaries of the sector, defined by two meridians, $\lambda = 0$ and $\lambda = \Xi$, where λ is longitude. The thick solid curves are the schematic isolines of a typical solution. Note that the two meridians meet at an angle, forming a corner at the north pole; a similar corner exists at the south pole (not visible).

purposes the simpler Helmholtz equation is better. For simplicity, we shall discuss only homogeneous Dirichlet boundary conditions in most of the article, but we shall explain the easy generalization to inhomogeneous boundary conditions in Section 6.

2. Corner singularities

Corner singularities are the theme of books by Grisvard [11] and Kozlow, Mazya, and Rossman [15] and Kondratiev's long review [14]; the numerical implications and various remedies have been discussed by many authors including [2,3,5,9,13,17,20,24,25]. For the present problem, note that (1) (for $k = 0$) has homogeneous solutions

$$\sin(s\lambda)P_n^s(\mu), \quad s = j(\pi/\Xi), \quad j = 1, 2, \ldots \tag{2}$$

It is well-known [1] that the associated Legendre function $P_n^s(\mu)$ is singular at both poles as $(1-\mu^2)^{s/2}$. The homogeneous and particular solutions for the Helmholtz equations are generally singular, too.

3. Mercator coordinate

Almost half a millenia ago, the prolific cartographer Gerhard Mercator introduced the stretched latitudinal coordinate that bears his name:

$$\mu = \tanh(y) \leftrightarrow y = \operatorname{arctanh}(\mu). \tag{3}$$

This transformation is useful because

$$(1 - \mu^2)^{s/2} = \operatorname{sech}^s(y). \tag{4}$$

The branch points at $\mu = \pm 1$ have been moved to infinity. As explained in [5], the hyperbolic secant function, raised to any power, decays exponentially as $|y| \to \infty$. Any reasonable basis set for the infinite interval will yield a spectral series that converges *exponentially fast* [8,5].

The chain rule implies

$$\frac{\mathrm{d}}{\mathrm{d}\mu} = \frac{1}{\operatorname{sech}^2(y)}\frac{\mathrm{d}}{\mathrm{d}y}, \quad (1-\mu^2)\frac{\mathrm{d}}{\mathrm{d}\mu} = \frac{\mathrm{d}}{\mathrm{d}y}. \tag{5}$$

The Helmholtz equation is transformed to

$$u_{\lambda\lambda} + u_{yy} + k^2 \operatorname{sech}^2(y)u = \operatorname{sech}^2(y) f(\lambda, \mu). \tag{6}$$

4. Chebyshev/rational Chebyshev pseudospectral method

Because the solution on a *sectorial* domain is not necessarily periodic in λ, the best spectral basis in longitude is Chebyshev polynomials. Because the standard interval for Chebyshev polynomials is $x \in [-1, 1]$, it is convenient to introduce the new longitudinal variable

$$x \equiv (2/\Xi)\lambda - 1 \rightarrow \frac{\mathrm{d}}{\mathrm{d}\lambda} = \frac{2}{\Xi}\frac{\mathrm{d}}{\mathrm{d}x}. \tag{7}$$

The Helmholtz equation becomes

$$\frac{4}{\Xi^2}u_{xx} + u_{yy} + k^2\operatorname{sech}^2(y)u = \operatorname{sech}^2(y)f((1+x)(\Xi/2), \mu), \quad x \in [-1, 1], \ \ y \in [-\infty, \infty]. \tag{8}$$

Our latitudinal basis is the set of rational Chebyshev functions [6], $TB_n(y; L)$. (Other choices are also possible as discussed in [4,5,16,22,23].) These have a width controlled by the constant L, the so-called "map parameter" that appears in their definition (9) below. This must be optimized by trial-and-error; $L = 2$ worked well for our examples.

Both sets of basis functions are images of a Fourier cosine basis under a change-of-coordinate:

$$T_m(\cos(t)) \equiv \cos(mt), \qquad TB_n(L\cot(s); L) = \cos(ns). \tag{9}$$

The basis functions and their derivatives can be calculated by using these trigonometric definitions, differentiating the cosines, and then transforming back to derivatives in x and y by applying the chain rule [8]

$$t = \arccos(x), \tag{10}$$

$$\frac{\mathrm{d}T_m}{\mathrm{d}x}(x) = m\frac{\sin(mt[x])}{\sin(t[x])}, \quad \frac{\mathrm{d}^2T_m}{\mathrm{d}x^2}(x) = -\frac{m^2}{\sin^2(t[x])}\cos(mt[x]) + \frac{m\cos(t[x])}{\sin^3(t[x])}\sin(mt[x]), \tag{11}$$

$$s(y) = \operatorname{arccot}(y/L), \tag{12}$$

$$\frac{\mathrm{d}TB_n}{\mathrm{d}y}(y) = \frac{n\sin^2(s[y])}{L}\sin(ns[y]), \tag{13}$$

$$\frac{\mathrm{d}^2TB_n}{\mathrm{d}y^2}(y) = \frac{1}{L^2}\left\{-n^2\sin^4(s[y])\cos(ns[y]) - 2n\cos(s[y])\sin^3(s[y])\sin(ns[y])\right\}. \tag{14}$$

One useful simplification is that the differential equation and boundary conditions are symmetric with respect to both $x = 0$ and $y = 0$. This implies that the computational task can be split into four subproblems. (Because the cost of factoring a dense matrix grows as the *cube* of the matrix size, it is cheaper by a factor of sixteen to solve four matrix problems of size $N_{\text{total}}/4$ than to solve a single matrix of size N_{total}.)

For simplicity, we shall describe the solution of the symmetric-symmetric subproblem. That is, we shall assume that $f(x, y) = f(-x, y) = f(x, -y)$ for all x and y and similarly for $u(x, y)$.

As explained in [8], an arbitrary function $g(x)$ can always be split into its symmetric and antisymmetric parts, $S(x)$ and $A(x)$, by $S(x) = (g(x) + g(-x))/2$ and $A(x) = (g(x)-g(-x))/2$. By applying this decomposition in both x and y, an arbitrary $f(x, y)$ can always be decomposed into the sum of a function f^{SS} which is symmetric in both x and y plus a function $f^{SA}(x,y)$ which is symmetric in x but antisymmetric in y plus a part which is antisymmetric in x but symmetric in y and finally a portion which is antisymmetric in both coordinates. Thus, we lose no generality by restricting attention to forcings and solutions that have definite symmetry in both coordinates.

To enforce the homogeneous boundary conditions, we write (for a function symmetric in both x and y)

$$u(x,y) = \sum_{m=1}^{M} \sum_{n=1}^{N} a_{mn} \{T_{2m}(x) - 1\}\{TB_{2n}(y;L) - 1\}. \tag{15}$$

Each basis function is individually zero at $x = \pm 1$, $y = \pm 1$ because $T_{2m}(\pm 1) = 1$ and $TB_{2n}(\pm 1) = 1$ for all m, n. The basis functions of odd degree are antisymmetric with respect to the origin, so to compute the part of the solution which is antisymmetric in x, we would replace $T_{2m}(x)-1$ by $T_{2m+1}(x)-x$, and similarly use $TB_{2n+1}(y;L)-TB_1(y;L)$ for parts of $u(x, y)$ which are antisymmetric with respect to the equator.

The pseudospectral method employs a grid which is all possible combinations (x_i, y_j) of the one-dimensional grids defined by

$$x_i = \cos\left(\frac{2i-1}{4M}\pi\right), \quad i = 1, 2, \ldots, M; \quad y_j = L\cot\left(\frac{2j-1}{4N}\pi\right), \quad j = 1, 2, \ldots, N. \tag{16}$$

Because of the double symmetry, the grid points are confined, for each of the four subproblems, to the upper right quadrant of the x–y plane, $x \in [0, 1]$ and $y \in [0, \infty]$.

Substituting the expansion into the differential equation and collocating at each of the MN grid points generates a linear $MN \times MN$ matrix problem which is solved by the usual LU factorization. When M, N are large, the $O([2/3]M^3N^3)$ cost of the factorization can be greatly reduced by solving the linear algebra problem using a preconditioned iteration instead of Crout reduction [8]. However, this is rarely necessary for a two-dimensional problem in a single unknown.

5. Numerical examples

5.1. Example one

As a test problem with a known exact solution, we chose

$$f(\lambda, \mu) = \cos(3\lambda)(1-\mu^2)^{3/2}(k^2 - 12), \quad u = \cos(3\lambda)(1-\mu^2)^{3/2} \tag{17}$$

with $k = 1$ and $\Xi = \pi/3$ so that the domain is one-sixth of the sphere. Note that u has a branch point at both poles, $\mu = \pm 1$. A double Chebyshev series would converge only as $O(1/N^3)$ [$O(1/N^4)$ for the coefficients], where N is the truncation in latitude ([8], p. 60). The exact solution is the spherical harmonic $Y_3^3(\lambda, \mu)$, so there is nothing atypical about this example.

The numerical results are illustrated in Fig. 2 and Table 1. An important point is that the rate of convergence with M, the number of Chebyshev polynomials, is geometric, that is, the error falls proportional to $\exp(-qM)$ for some constant q. However, it is known [6,8] that the normal rate of convergence for expansions on an *infinite* interval is "subgeometric", that is, proportional to an exponential whose argument is a fractional power of the truncation, typically $\exp(-q'\sqrt{N})$, where q' is another constant. Thus, we expect that the rate of convergence will be *highly anisotropic*: it is necessary to choose N (truncation in the TB series in y) much larger than M.

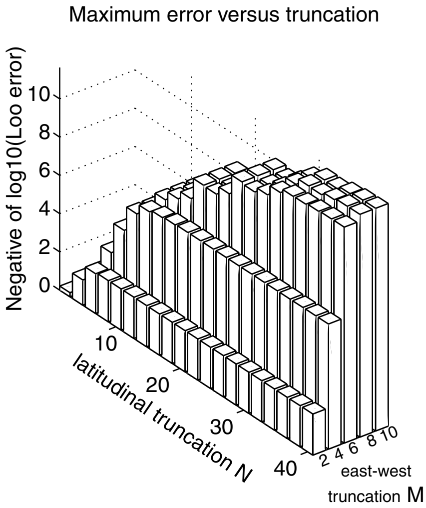

Fig. 2. The negative of the base-10 logarithm of maximum pointwise (L_∞) error versus the truncations M and N. Thus, a vertical height of 8 is equivalent to an absolute error of 10^{-8}.

Table 1
Errors versus resolution for the test problem with $L = 2$

M	N	Maximum pointwise error
4	4	0.04
4	8	4.5E−5
8	8	3.4E−5
8	16	2.8E−8
8	32	4.1E−12
8	64	2.2E−12

Fig. 2 shows that when the longitudinal truncation M is small, such as $M = 2$ or $M = 4$, the error saturates rapidly as N increases: there is no point in using a lot of north–south basis functions when the east–west resolution is too small. However, when $M = 6$ or larger, the error decreases (i.e., the bars in the diagram grow *taller*) until $N = 30$. At this point the error is $O(10^{-12})$, far below any reasonable scientific/engineering requirements!

Because of roundoff, the error does not decrease further with increases in M and N beyond $M = 6$, $N = 30$. Machine epsilon is 2×10^{-16} for our system, but the error in spectral calculations usually "plateaus" at a level which is a hundred to ten thousand times greater as explained in [8].

One possible snag with the Mercator coordinate is that the exponentials in the mapping can lead to underflow and numerical ill-conditioning when N is large. These can be mitigated by varying the map parameter L and such tricks as replacing $(1-\tanh(y))$ by its asymptotic approximation. However, these refinements were quite unnecessary for our examples.

5.2. Example two

We also solved the Helmholtz equation with the forcing

$$f = \sin(3\lambda)(1-\mu^2)^{3/2}\frac{1}{[36/\pi^2](\lambda-\pi/6)^2+1/10}\,\frac{1}{\mu^2+1/10}. \tag{18}$$

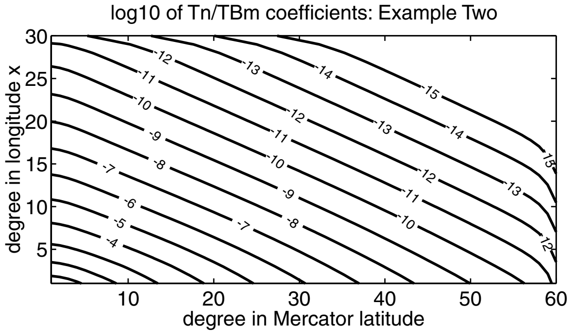

Fig. 3. Isolines of the base-10 logarithm of the absolute values of the spectral coefficients a_{mn} plotted versus latitudinal and longitudinal degree. The largest coefficient is 0.63.

No exact solution is known, but this example has poles in the complex plane along both the imaginary μ axis (at $\mu = \pm i\sqrt{1/10}$) and imaginary x axis ($x = \pm i\sqrt{1/10}$) as well as branch points at both the north and south poles. The rational factors were chosen to be equally demanding in λ and μ so that any observed anisotropy would reflect T_n/TB_n differences rather than differences in the scale of variation of the solution in the two different coordinates. Because of the singularities, more basis functions are needed to achieve high accuracy and the anisotropy is less pronounced: one needs roughly twice as many latitudinal basis functions as longitudinal basis functions to achieve a given accuracy. Fig. 3 shows the coefficients: it is remarkable that all coefficients near the truncation limits $M = 30$, $N = 60$ are $O(10^{-11})$ or smaller!

6. Generalizations

6.1. Inhomogeneous boundary conditions

If the boundary conditions are

$$u(0,\mu) = \phi_{\text{left}}(\mu), \quad u(\Xi,\mu) = \phi_{\text{right}}(\mu), \tag{19}$$

the problem can be transformed to one with homogeneous boundary conditions for a new variable v by writing

$$u \equiv v(\lambda,\mu) + g(\lambda,\mu), \tag{20}$$

where

$$g(\lambda,\mu) \equiv \phi_{\text{left}}(\mu) + (\lambda/\Xi)\{\phi_{\text{right}}(\mu) - \phi_{\text{left}}(\mu)\}. \tag{21}$$

Since g satisfies the desired boundary conditions, $v = 0$ on both meridians. It solves the Helmholtz equation with the modified forcing

$$\nabla^2 v + k^2 v = f - \nabla^2 g - k^2 g, \tag{22}$$

where ∇^2 is the surface Laplace operator written out explicitly in (1).

Alternatively, the basis functions can deal with inhomogeneous boundary conditions directly [8].

6.2. Four corners

In a rectangular domain, the Mercator change-of-coordinate can be applied in *both* coordinates to map the four corners to infinity and render them harmless. The spectral basis is a tensor product of rational Chebyshev functions.

7. Alternative: separation-of-variables

The Helmholtz equation in a sector can also be solved by the classic eighteenth century method of separation-of-variables. Expand

$$u = \sum_{m=1}^{\infty} \sum_{n}^{\infty} u_{mn} \sin(mS\lambda) P_{mS+n}^{-mS}(\mu), \tag{23}$$

where $S = \pi/\Xi$ so that the boundary conditions are satisfied and the associated Legendre functions are

$$P_{s+n}^{-s}(\mu) \propto (1-\mu^2)^{s/2} C_n^{s+1/2}(\mu), \tag{24}$$

where the $C_n^{s+1/2}$ is the nth degree Gegenbauer polynomial. By invoking the differential equation satisfied by P_v^s, which is

$$(1-\mu^2)\frac{d^2 P_v^s}{d\mu^2} - 2\mu \frac{dP_v^s}{d\mu} - \frac{s^2}{1-\mu^2} P_v^s = -v(v+1)P_v^s \tag{25}$$

and similarly expanding the inhomogeneous term in the Helmholtz equation with coefficients f_{mn}, one finds

$$b_{mn} = \frac{f_{mn}}{k^2 - (mS+n)(mS+n+1)}. \tag{26}$$

The coefficients of f_{mn} can be computed by numerical evaluation of a double integral (in λ and μ) and the Gegenbauer polynomials can be evaluated by three-term recurrence relation. It all seems so fast and easy that one wonders why anyone but a blockhead would use anything else.

There are a couple of reasons. First, the Chebyshev/rational Chebyshev program is very short with only seventy-one statements (excluding graphical output). It would be difficult to write a program to compute and evaluate the Fourier–Legendre series with fewer lines; the associated Legendre/Gegenbauer functions require some investment of both learning time and programming time. The quadratures are tricky because the integrands have branch points at both ends of the interval of integration. L.N. Trefethen's maxim is applicable here: "Just because there's an exact formula doesn't mean it's necessarily a good idea to use it".

Second, because the forcing and solution are usually not periodic in longitude when the domain is only a sector, the Fourier series will converge at only an algebraic rate – that is, the error will decrease as a finite inverse power of M instead of exponentially with M, where M is again the truncation of the longitudinal series. If $f(\lambda, \mu)$ does not equal zero on both boundaries, then its series will exhibit the Gibbs' Phenomenon [8] and its coefficient will decrease proportional to $1/M$, implying that the Fourier coefficients of $u(\lambda, \mu)$ will decrease only as $O(1/M^3)$.

This is in fact a generic difficulty of separation-of-variables series solutions, never discussed in classic texts such as [12,18]. Unless one is content with an answer of very modest accuracy, the Chebyshev/rational Chebyshev expansion is the better way.

8. Conclusions

There is nothing radically novel in this article in the sense that the use of a Mercator coordinate to defeat corner singularities has been suggested before [5,22]. However, the sector-of-sphere problem is interesting because of its anisotropy: there are only two corners instead of four as in a rectangle, and the Mercator transformation is applied in only one coordinate instead of all. Does this anisotropy cause difficulties? Does the combination of a Chebyshev basis without transformation with a rational Chebyshev basis in a transformed coordinate degrade the spectral method?

The answer to both these questions is a resounding No! However, the double series is highly anisotropic in the sense that, all other things being equal, one needs significantly more basis functions in latitude than in longitude. The rate of convergence is geometric in x but only subgeometric in y. One needs roughly five times as many basis functions in y as in longitude for our first example and a ratio of about two to one for our second.

The method of separation-of-variables is unsatisfactory even though it is usually regarded as an analytic rather than a numerical method: it yields slowly-converging series whose coefficients decrease as inverse powers of the truncations M and N. In contrast, the Chebyshev/rational Chebyshev numerical method has an *exponential* rate of convergence.

Acknowledgements

This work was supported by the National Science Foundation through Grant OCE9986368. I thank the reviewers for helpful comments.

References

[1] M. Abramowitz, I.A. Stegun, Handbook of Mathematical Functions, Dover, New York, 1965.
[2] M. Amara, C. Bernardi, M.A. Moussaoui, Handling corner singularities by mortar elements method, Appl. Anal. 46 (1992) 25–44.
[3] A. Averbuch, M. Israeli, L. Vozovoi, A fast Poisson solver of arbitrary order accuracy in rectangular regions, SIAM J. Sci. Comput. 19 (1998) 933–952.
[4] K.L. Bowers, J. Lund, Numerical solution of singular Poisson equations by the sinc-Galerkin method, SIAM J. Numer. Anal. 24 (1987) 36–51.
[5] J.P. Boyd, Polynomial series versus sinc expansions for functions with corner or endpoint singularities, J. Comput. Phys. 64 (1986) 266–269.
[6] J.P. Boyd, Spectral methods using rational basis functions on an infinite interval, J. Comput. Phys. 69 (1987) 112–142.
[7] J.P. Boyd, Traps and snares in eigenvalue calculations with application to pseudospectral computations of ocean tides in a basin bounded by meridians, J. Comput. Phys. 126 (1996) 11–20, Corrigendum 136(1) (1997) 227–228.
[8] J.P. Boyd, Chebyshev and Fourier Spectral Methods, 2nd ed., Dover, Mineola, New York, 2001, p. 665.
[9] E. Braverman, M. Israeli, A. Averbuch, L. Vozovoi, A fast 3D Poisson solver of arbitrary order accuracy, J. Comput. Phys. 144 (1998) 109–136.
[10] C. Sozou, On some forced oscillations relating to Laplace's tidal equation, Geophys. J. Int. 119 (1994) 779–782.
[11] P. Grisvard, Elliptic Problems in Nonsmooth Domains, Pitman, New York, 1985.
[12] J.D. Jackson, Classical Electrodynamics, John Wiley and Sons, New York, 1962.
[13] M. Kermode, A. McKerrell, L.M. Delves, The calculation of singular coefficients, Comput. Meths Appl. Mech. Eng. 50 (1985) 205–215.
[14] V.A. Kondratiev, Boundary value problems for elliptic equations in domain with conical or angular points, Trans. Moscow Math. Soc. 16 (1967) 227–313.
[15] V.A. Kozlov, V.G. Mazya, J. Rossman, Spectral problems associated with corner singularities of solutions of elliptic equations, Mathematical Surveys and Monographs, 85, American Mathematical Society, Providence, RI, 2000.
[16] J. Lund, K.L. Bowers, Sinc Methods For Quadrature and Differential Equations, Society for Industrial and Applied Mathematics, Philadelphia, 1992, p. 304.
[17] A. McKerrell, The global element applied to fluid flow problems, Comput. Fluids 16 (1988) 41–46.
[18] P.M. Morse, H. Feshbach, Methods of Theoretical Physics, McGraw-Hill, New York, 1953, p.2000.
[19] W.P. O'Connor, The complex wavenumber eigenvalues of Laplace's tidal equations for ocean bounded by meridians, Proc. R. Soc. Lond. A 449 (1995) 51–64, Corrigendum A 452 (1996) 1185–1187.
[20] W.W. Schultz, N.-Y. Lee, J.P. Boyd, Chebyshev pseudospectral method of viscous flows with corner singularities, J. Scientific Comput. 4 (1989) 1–24.
[21] C. Sozou, On azimuthal eigenwavenumbers associated with Laplace's tidal equation, Geophys. J. Int. 130 (1997) 151–156.
[22] F. Stenger, Sinc Methods, Springer-Verlag, New York, 1993, p.500.

[23] F. Stenger, Summary of sinc numerical methods, J. Comput. Appl. Math. 121 (2000) 379–420.
[24] C. Xenophontos, The hp finite element method for singularly perturbed problems in nonsmooth domains, Numer. Meths. Partial Diff. Eqs. 15 (1999) 65–89.
[25] Z. Yosibash, B. Szabo, Numerical analysis of singularities in two dimensions: Part 1: Computation of eigenpairs, Int. J. Numer. Meths. Eng. 38 (1995) 2055.