A simple algorithm for the inverse field of values problem

Russell Carden

Department of Computational and Applied Mathematics, Rice University, Houston, TX 77005-1832, USA

E-mail: Russell.L.Carden@rice.edu

Abstract.

The field of values of a matrix is the closed convex subset of the complex plane comprising all Rayleigh quotients, a set of interest in the stability analysis of dynamical systems and convergence theory of matrix iterations, among other applications. Recently, Uhlig proposed the inverse field of values problem: given a point in the field of values, determine a vector for which this point is the corresponding Rayleigh quotient. Uhlig also devised a sophisticated algorithm involving random vectors and the boundaries of ellipses for solving the inverse field of values problem. We propose a simpler deterministic algorithm that must converge (in exact arithmetic), and for most points yields an exact result in only a few iterations. The algorithm builds upon the fact that the inverse field of values problem can be solved exactly in the two dimensional case. We also resolve a conjecture posed by Uhlig concerning the number of linearly independent vectors that generate a point in the field of values, and propose a more challenging inverse field of values problem that is of interest in eigenvalue computations.

AMS classification scheme numbers: 65F35,15A29,15A60

Submitted to: Inverse Problems

1. Introduction

The field of values (also known as the numerical range) of a matrix $A \in \mathbb{C}^{n \times n}$ is a closed, convex subset of the complex plane defined as

$$W(A) \equiv \{x^*Ax : x \in \mathbb{C}^n, ||x|| = 1\},$$

where x^* denotes the conjugate transpose of x and $||x||^2 = x^*x$. Hence, the field of values is the set of all Rayleigh quotients for a matrix A. Given some $z \in W(A)$, we call a unit vector x such that $z = x^*Ax$ a generating vector for z. In the context of iterative eigensolvers, the point z and a corresponding generating vector x are referred to as a Ritz value/Ritz vector pair. Quantities derived from the field of values, such as the numerical radius and the numerical abscissa, provide insight into the behavior of matrix

powers and matrix exponentials; see, e.g., [5, 12]. In the restarted Arnoldi method for computing eigenvalues [11], Ritz values serve not only as eigenvalue estimates, but also as the roots of restart polynomials used to generate vectors that yield better eigenpair approximations. Understanding the convergence of such iterative eigensolvers requires insight into how Ritz values distribute themselves throughout the field of values [1, 2].

A preliminary step toward understanding how Ritz values must distribute themselves involves the inverse field of values problem recently proposed by Uhlig [13]: given a point $z \in W(A)$, determine a unit vector x such that $z = x^*Ax$. To solve the inverse field of values problems, Uhlig proposed an algorithm that begins by generating points in the field of values from random unit vectors, in an attempt to surround the desired z. Uhlig's algorithm then iteratively refines these preliminary generating vectors to better approximate a generating vector for z. Due to the difficulty of encircling the desired point using Rayleigh quotients associated with random vectors, the algorithm does not always converge to the desired accuracy, see figure 4. In posing the inverse field of values problem, Uhlig hoped to derive some insight into the structure of the field of values using the notion of the covering number, the number of linearly independent generating vectors for a particular point in the field of values. Uhlig conjectured that the covering number of any point in the interior of the field of values must be n.

We present a simple algorithm for solving the inverse field of values problem. Our algorithm is comparable to the standard algorithm used to determine the boundary of the field of values of a matrix [9]. We prove that indeed the covering number is n for any point in the interior of the field of values. Finally we propose a more general inverse field of values problem that involves the construction of subspaces of dimension greater than one that yield Rayleigh quotients corresponding to several points in W(A). The solution to this rather more challenging problem would provide considerable insight into how Ritz values distribute themselves throughout the field of values.

2. Main Ideas

In this section we state two results that are essential for the algorithm that follows. We present solutions to the inverse field of values problem for two scenarios: (1) two dimensional matrices; (2) the point z is in the convex hull of three Ritz values, θ_i , for which we have the corresponding generating vectors, x_i . In both scenarios we exactly determine a unit vector x such that $z = x^*Ax$. Both results follow directly from the argument used to prove that the field of values is convex [7, Ch.1].

The solution for the first scenario follows from the proof of Lemma 1.3.1 in [7]. Without loss of generality, one may assume that the trace of A is zero. (If the trace were nonzero one would need only subtract $(\operatorname{tr}(A)/2)I$ from A and $\operatorname{tr}(A)/2$ from z.) With this assumption the eigenvalues sum to zero, and may be denoted as $+\lambda$ and $-\lambda$. Furthermore, one may assume that eigenvalues of A are real. (To accomplish this, one need only multiply both A and z by $e^{-i\psi}$ where $\lambda = |\lambda|e^{i\psi}$.) As in [7, §1.3], one may

determine a unitary matrix Q such that

$$Q^*AQ = \widehat{A} = \begin{bmatrix} 0 & \alpha \\ \beta & 0 \end{bmatrix},$$

where α and β are both real and nonnegative and $\alpha \geq \beta$. Note that z need not be altered, as the field of values is invariant under unitary similarity transformations. As any 2×2 matrix can be shifted, scaled and unitarily transformed into this form, if we can solve the inverse field of values problem for \widehat{A} , then we are done.

Without loss of generality, we can represent the unit vector $x \in \mathbb{C}^2$ as $x = [\cos(\theta); e^{i\phi}\sin(\theta)]$, giving

$$x^* \widehat{A} x = \frac{\sin(2\theta)}{2} ((\alpha + \beta)\cos(\phi) + (\alpha - \beta)i\sin(\phi)),$$

from which one can see that the field of values of \widehat{A} is an ellipse in the complex plane centered at the origin with major axis $\alpha + \beta$ and minor axis $\alpha - \beta$. Given $z \in W(\widehat{A})$, the θ and ϕ that specify the generating vector x must satisfy

$$\tan(\phi) = \frac{(\alpha + \beta)\operatorname{Im}(z)}{(\alpha - \beta)\operatorname{Re}(z)}, \qquad \sin(2\theta) = \frac{2|z|}{|(\alpha + \beta)\cos(\phi) + (\alpha - \beta)\sin(\phi)|}.$$

In terms of the coordinates for A, the generating vector is Qx. Thus we have solved the inverse field of values problem for two-dimensional matrices. Note that the solution need not be unique, i.e., for z in the interior of W(A), there exist linearly independent generating vectors, θ and $\pi/2 - \theta$ for $\theta \neq \pi/4$. (When $\theta = \pi/4$, z is a point on the boundary, and its generating vector is unique provided that the ellipse is nondegenerate. If $\alpha = \beta$, then the field of values of A is a line segment connecting the eigenvalues, and the solution is unique only for the end-points, which must also be eigenvalues. If $\alpha = \beta = 0$, then A is the zero matrix and the field of values is just a point and any unit vector $x \in \mathbb{C}^2$ will do.)

For the second scenario, we first recall why convexity of the field of values for 2×2 matrices implies the same for all larger matrices. Given any two unique points in the field of values, θ_1 and θ_2 , and their corresponding generating vectors, x_1 and x_2 , one can construct a matrix $U \in \mathbb{C}^{n \times 2}$ with orthonormal columns such that $x_1, x_2 \in \text{Range}(U)$. Then the field of values of the restriction of A to Range(U), $H = U^*AU$, must be contained in the field of values of A. By construction, the field of values of H must contain both θ_1 and θ_2 . Since the field of values of a 2×2 matrix is an ellipse and hence convex, W(H) as well as W(A) must contain all convex combinations of θ_1 and θ_2 . Since θ_1 and θ_2 were arbitrary, W(A) must be convex. Thus the solution to our first scenario indicates that if we are given two values θ_1 and θ_2 in the field of values of A, then we can determine vectors that generate any point that is a convex combination of θ_1 and θ_2 .

Now suppose z is in the convex hull of three points, $\theta_i \in W(A)$, for which we have the corresponding generating vectors x_i . The convex hull of the three points is a (possibly degenerate) triangle. To apply the result for the 2×2 case, we would need z to fall on a segment with end-points for which we have — or can compute — corresponding

generating vectors. Without loss of generality, take θ_1 as one of the end-points for this proposed line segment. For the other end-point of the line segment, take the point w to be the intersection of the line through θ_2 and θ_3 with the line through θ_1 and z. As w is a convex combination of θ_2 and θ_3 , we can determine a corresponding generating vector for w. Now since z is a convex combination of w and θ_1 , we can determine a generating vector for z, as illustrated in figure 1.

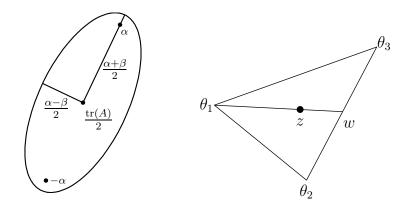


Figure 1. On the left, an ellipse corresponding to the field of values of a general 2×2 matrix. On the right, an illustration of the procedure for determining a generating vector for a point located in the convex hull of three points in W(A).

Given the results above, all that remains is to specify a way to generate a triangle in W(A) that contains the desired value z. For this we consider Johnson's algorithm for the numerical determination of the field of values of a general complex matrix [7, 9]. Johnson's method constructs an approximation to the field of values by determining the eigenvalues of $H_{\theta} = (e^{i\theta}A + e^{-i\theta}A^*)/2$, the Hermitian part of $e^{i\theta}A$, for various values of θ . The largest and smallest eigenvalues of H_{θ} provide sharp bounds on the extent of the field of values in the direction θ in the complex plane. These sharp bounds together form an outer polygonal approximation to the field of values, i.e., a polygon that contains W(A) and whose sides intersect points on the boundary of W(A). This outer approximation can be used to identify points that fall outside the field of values. The corresponding eigenvectors of H_{θ} generate points on the boundary of W(A). These points, taken together, yield an inner polygonal approximation to W(A); see figure 3. By systematically selecting a series of angles θ , one can construct polygons that, to a desired accuracy, contain the desired value z. This is the substance of the algorithm we describe in Section 4.

3. Covering number

Before detailing our algorithm, we pause to resolve a question proposed by Uhlig [13].

Definition 1. For a matrix $A \in \mathbb{C}^{n \times n}$ and a point $z \in W(A) \subset \mathbb{C}$ we call the maximal

dimension of a linearly independent set of complex unit vectors x with $x^*Ax = z$ the covering number of (A, z).

Having shown that the generating vectors for a point in the interior of W(A) for $A \in \mathbb{C}^{2\times 2}$ is not unique, we have demonstrated that in the 2×2 case, the covering number of any point in the interior of W(A) is 2.

Uhlig introduced the covering number as a tool that might provide insight into the structure of the field of values, particularly for nonnormal matrices; however, based on numerical observations he conjectured that the covering number is n for all points in the interior of the field of values, which would imply that this initially appealing idea yields no further insight into the structure of the field of values. We can confirm that his conjecture holds.

Theorem 1. Let $A \in \mathbb{C}^{n \times n}$ and suppose z is in the interior of W(A). Then the covering number of (A, z) is n.

Proof. Consider any point z in the interior of W(A). (Note that if W(A) is a line segment, then by interior we mean that z is not an end-point of the line segment.) By definition, all points in W(A) have at least one generating vector. Suppose that we have a set of k < n linearly independent generating vectors for z. We will show that we can always find another linearly independent generating vector. As k < n, there exists a unit vector u that is orthogonal to the set of k linearly independent generating vectors for z. Consider the point $a \in W(A)$ generated by u, noting that $a \neq z$. Since z is in the interior of W(A) and W(A) is convex, there always exists some $b \in W(A)$, $b \neq z, a$, for which z is a convex combination of a and b. Since $b \in W(A)$, it must have some generating vector, \hat{u} . Now z must be in the interior of the field of values of the 2×2 matrix corresponding to the restriction of A onto the span $\{u, \widehat{u}\}$. Thus from our argument in Section 2, we find that there exist two linearly independent generating vectors for z such that u is in their span. At least one of these vectors must have a component in u, thus giving an additional generating vector for z linearly independent from the rest. Hence we have at least k+1 linearly independent generating vectors for z. As k < n and z were arbitrary, the covering number for any point in the interior of the field of values must be n.

4. Algorithm

Using the results above one can specify an algorithm for solving the inverse field of values problem. Given a matrix $A \in \mathbb{C}^{n \times n}$ and some $z \in \mathbb{C}$, let $\varepsilon > 0$ denote some tolerance (for example $\varepsilon = 10^{-16} ||A||$ for a double precision computation).

I. Determine the leftmost and rightmost eigenvalues of H_{θ} for $\theta = 0, \pi/2$. The resulting outer approximation will be a rectangle with sides parallel to the real and imaginary axes. If z is inside the outer approximation, then continue; otherwise conclude $z \notin W(A)$.

- II. If the height or width of the outer approximation is less than ε , then W(A) is approximately either Hermitian or skew-Hermitian (or a complex shift of one of these). If both the height and width of the outer approximation are less than ε , then the field of values is approximately a point. In either of these cases, one can easily determine if z is in W(A) and if so determine a corresponding generating vector; otherwise, continue.
- III. Using the eigenvectors of the leftmost and rightmost eigenvalues of H_{θ} for $\theta = 0, \pi/2$, determine an inner approximation to W(A). The inner approximation will be a quadrilateral with vertices on the boundary of W(A).
- IV. Check if z lies in the inner approximation. If so, determine a vector that exactly generates z as described in Section 2, and successfully terminate. Otherwise, determine the side of the inner approximation to which z is closest.
- V. Compute \hat{z} , the closest point to z from the inner approximation. If $|\hat{z} z| < \varepsilon$ then compute the corresponding generating vector for \hat{z} , accept this vector as a suitable generating vector for z, and successfully terminate.
- VI. Update the inner and outer approximations by determining the largest eigenvalue and corresponding eigenvector of H_{θ} , where θ is the direction perpendicular to the side of the inner approximation to which z is closest. If this does not produce a new boundary point not previously contained in the inner approximation, then $z \notin W(A)$.
- VII. Check if z is in the new outer approximation. If so, go to step IV; otherwise, $z \notin W(A)$.

An illustration of the algorithm is shown in figures 2 and 3.

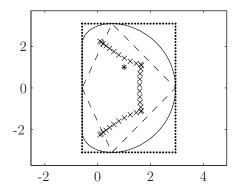


Figure 2. Eigenvalues (\times), field of values boundary (solid line) and first inner and outer polygonal approximations (dashed, dotted) for Matlab's Grear matrix (gallery('grear',32)). In this case z (denoted by the *) falls inside he inner approximation, and an exact generating vector can be determined immediately.

Note that steps IV–VII are repeated until the inner approximation is within ε of z, or until the outer approximation does not contain z. This algorithm only updates the approximation to the field of values in the direction of z. For determining if z lies

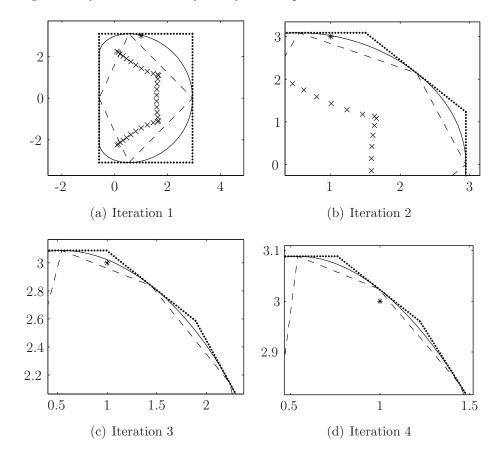


Figure 3. Eigenvalues (\times), field of values boundary (solid line) and first inner and outer polygonal approximations (dashed, dotted) for four iterations of the algorithm for Matlab's Grear matrix (gallery('grear',32)) for z=1+3i. At the fourth iteration, z (denoted by *) is in the inner approximation and a generating vector that exactly gives z can be determined.

inside a particular polygonal approximation, there exist many efficient algorithms; see, e.g., [4]. In most cases the algorithm will terminate exactly in step IV after only a few eigenanalyses.

The algorithm above does not utilize the fact that the field of values of a 2×2 matrix is an ellipse. This property suggests that with the points and the corresponding generating vectors for the inner approximation, one may be able to exactly determine generating vectors for some points outside the inner approximation. This fact can be utilized with a modification of step V.

V(a) Check if z is in the ellipse corresponding to the field of values of the restriction of A to the subspace spanned by the two generating vectors associated with the end points of the side of the inner approximation closest to z. If so, determine a vector that exactly generates z. Otherwise, compute \hat{z} , the closest point to z from the ellipse. If $|\hat{z} - z| < \varepsilon$, then compute the corresponding generating vector for \hat{z} and accept it as a suitable generating vector for z, and successfully terminate.

Determining the closest point from an ellipse, a problem first solved by Apollonius, involves solving a quartic polynomial [6]. Depending on the nature of the ellipse, this modification could reduce the number of iterations required by one.

5. Discussion

The proposed algorithm differs from Uhlig's in several ways. Our algorithm is nearly deterministic; the only lack of determinism occurs if the extreme eigenvalues of H_{θ} have multiplicity greater than one, which is only possible when the boundary of the field of values contains a line segment. Otherwise, for a given z the algorithm will always return the same generating vector. A method for finding other generating vectors is suggested in the proof of Theorem 1. At the cost of determinism, Uhlig attempts to minimize the number of eigenanalyses required by using random vectors to generate points in the field of values. If the randomly generated points sufficiently surround the desired point, then Uhlig's algorithm refines a quadrilateral that contains the desired point. Using great unit circles in \mathbb{C}^n that contain pairwise generating vectors associated with the vertices of the quadrilateral, he refines his quadrilateral using the images of these great circles under the map $x \mapsto x^*Ax$, which happen to be ellipses. Points on these ellipses are determined numerically using equally spaced vectors on the great circle. We have demonstrated how one could use the generating vectors for the quadrilateral to exactly solve the problem, rather than iterates as in Uhlig's approach. We utilize ellipses associated with subspaces spanned by pairs of generating vectors. For these ellipses we can exactly generate any point in the interior as well as on the boundary; thus we are able to determine generating vectors for any point in the quadrilateral. Were Uhlig's algorithm to take full advantage of these ellipses, it could potentially be faster and more accurate, particularly for points not near the boundary of the field of values.

As the Rayleigh quotients associated with random vectors tend to lie in the interior of W(A), Uhlig's algorithm has difficulty generating Rayleigh quotients that surround points near the boundary of W(A). Also, as Uhlig observed, this problem becomes more acute for higher dimensions. If the randomly generated points do not sufficiently contain the desired point, Uhlig's algorithm utilizes the eigenvectors of H_{θ} for no more than six distinct θ . The θ are chosen in much the same way that our algorithm determines a polygon that contains the desired point. Uhlig's concern in limiting the number of eigenanalyses that must be performed is that each eigenanalysis requires $O(n^3)$ computations, whereas the refinement process only requires matrix vector operations of complexity $O(n^2)$. Our algorithm performs a minimum of two eigenanlyses and could potentially perform more than six eigenanalyses, depending upon the nature of the boundary of W(A). To address this concern, in our implementation of the algorithm we allow for the use of Matlab's sparse eigensolver, eigs (which uses the ARPACK software [10]). In this case, for $\theta = 0, \pi/2$, a total of four calls must be made to eigs to determine just the extreme eigenvalues of the H_{θ} . For additional angles utilized in step VI, only one call to eigs is necessary, as only the largest real eigenvalue is required.

The advantage of using random vectors versus eigenvalue computations to generate a quadrilateral that contains the desired point will vary with the size of the matrix and the distance of z from the boundary of W(A). Even with the concern for the number of eigenanalyses required, our algorithm, using Matlab's dense eigensolver eig, often performed much faster for points throughout the majority of W(A) for A of a variety of dimensions. Implementations of Uhlig's method and our algorithm are publicly available [3, 14].

A comparison of run-times and errors for a Jordan block and a circulant shift,

$$\begin{pmatrix} 0 & 1 & & & \\ & \ddots & \ddots & & \\ & & 0 & 1 \\ & & & 0 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} 0 & 1 & & \\ & \ddots & \ddots & \\ & & 0 & 1 \\ 1 & & & 0 \end{pmatrix},$$

are shown in figure 4. These matrices were chosen to illustrate extremes: a matrix with a large departure from normality, and a normal matrix for which the boundary of W(A) contains line segments. For the Jordan block, one can see that with the exception of points on the boundary, our algorithm is nearly an order of magnitude faster for points throughout the field of values. For the circulant matrix, our algorithm clearly performs faster for all points not near the boundary. The error plot in figure 4(c) shows that for Uhlig's algorithm, the error (the difference between the Rayleigh quotient from the computed generating vector and the desired z) varies over ten orders of magnitude throughout the field of values, with the majority of the errors of the order 10^{-12} . The error for our algorithm as shown in figure 4(d) is much more uniform, varying only an order of magnitude about 10^{-14} . These results are typical of a range of test cases we have investigated.

6. A more general inverse problem

We propose an inverse field of values problem, iFOVk, involving multiple Ritz values:

Given
$$z_1, \ldots, z_k \in W(A)$$
, does there exist $U \in \mathbb{C}^{n \times k}$, $U^*U = I$ such that the spectrum of U^*AU is $\{z_1, \ldots, z_k\}$ and if so, find such a U .

Note that the z_i need not be distinct, e.g, $z_1 = z_2$. The algorithm in this paper is concerned with iFOV1. For iFOVn, the only possible set $\{z_i\}$ is the spectrum of A. Even when k = 2, there exist matrices A and pairs $\{z_1, z_2\} \subset W(A)$ so that no such U exists (e.g., the Courant–Fischer min-max theorem [8]). The solution to this problem for general k would allow one to study how Ritz values distribute themselves throughout the field of values. Such knowledge would prove useful for analyzing eigenvalue algorithms such as the restarted Arnoldi method, and could suggest a kind of generalization of the interlacing theorem for non-Hermitian matrices.

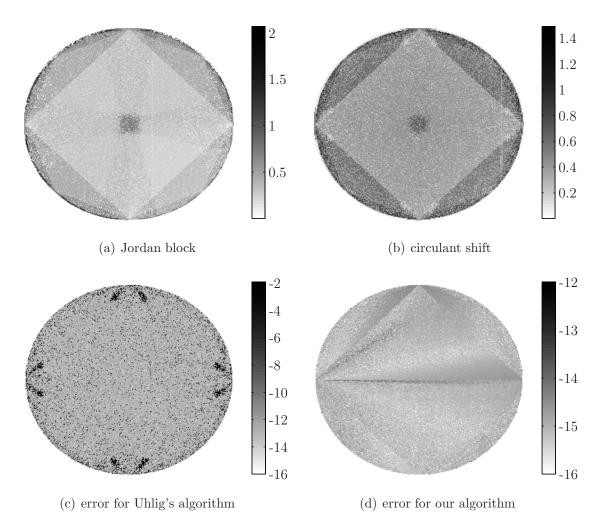


Figure 4. Comparison of Uhlig's algorithm versus our method for a Jordan block and a circulant shift matrix of dimension n=240 for those points on a 256×256 uniform grid for $[-1,1]\times [-1,1]$ that fall in W(A). (a) &(b) Ratio of run time for our method to run time for Uhlig's algorithm, (c)&(d) \log_{10} of error for circulant matrix.

Acknowledgments

None of this work would have been done had not Mark Embree pointed me to Uhlig's paper. Embree provided vital feedback on various versions of the algorithm as well as revisions of this paper.

References

- [1] Beattie C A, Embree M and Sorensen D C 2005 Convergence of polynomial restart Krylov methods for eigenvalue computations SIAM Review 47 p 492-515
- [2] Carden R 2009 Ritz values and Arnoldi convergence for nonsymmetric matrices, Masters Thesis (Rice University)
- [3] Carden R 2009 MATLAB m-file inversefov.m (available at http://www.caam.rice.edu/tech_reports/2009/inversefov.m)

- [4] Cormen T H, Leiserson C E, Rivest R L and Stein C 2001 Introduction to Algorithms, Second Edition (Cambridge: The MIT Press) p 889-890
- [5] Crouiex M 2005 Numerical range and functional calculus in Hilbert space Journal of Functional Analysis 244 p 668-690
- [6] Hartmann F and Jantzen R. Apollonius's Ellipse and Evolute Revisited *Loci*. Retrieved June 4, 2009 from http://mathdl.maa.org/mathDL/46/?pa=content&sa=viewDocument&nodeId=2554
- [7] Horn R A and Johnson C R 1991 Topics in Matrix Analysis (Cambridge: Cambridge University Press) p 1-64
- [8] Horn R A and Johnson C R 1985 Matrix Analysis (Cambridge: Cambridge University Press) p 179-180
- [9] Johnson C R 1978 Numerical determination of the field of values of a general complex matrix $SIAM\ J.\ Num.\ Anal.\ {\bf 15}$ p 595-602
- [10] Lehoucq R B, Sorensen D C and Yang C 1998 ARPACK user's guide (Philadelpia:SIAM)
- [11] Sorensen D C 1992 Implicit application of polynomial filters in a k-step Arnoldi method SIAM J. Matrix Anal. Appl. 13 p 357-385
- [12] Trefethen L N and Embree M 2005 Spectra and Pseudospectra: The Behavior of Nonnormal Matrices and Operators (Princeton:Princeton University Press) p 166-175
- [13] Uhlig F 2008 An inverse field of values problem Inverse Problems 24 055019
- [14] Uhlig F 2008 MATLAB m-file wberpoint.m (available at http://www.auburn.edu/~uhligfd/m_files/wberpoint.m)