Chapter I

Introduction

1 Background and Scope

A prototype of an eigenvalue problem is to determine nontrivial solutions of the equation

$$Ax = \lambda x, \tag{1.1}$$

where $A \in \mathbb{C}^{n \times n}$ is given, and $x \neq 0 \in \mathbb{C}^n$ and $\lambda \in \mathbb{C}$ are unknown. The value λ is called an *eigenvalue* and x is a corresponding *eigenvector*.

In general, equation (1.1) has more than one solution. The eigenvalues are also zeros of the polynomial equation

$$\det(A - \lambda I) = 0, \tag{1.2}$$

and this equation has n roots. To each different root there corresponds a different eigenvector, but for multiple roots there may be only one single eigenvector solution. If equation (1.1) has less than n independent solutions for x, then the matrix is called *defective*.

In this book we will see different types of eigenvalue problems. An often used approach to solve these other problems is to transform them in some convenient way to the standard eigenvalue problem (1.1).

Eigenvalue problems have been important since the early days of scientific computing. Indeed, in 1846, Jacobi [1846] published a paper that describes how to compute eigenvalues for small linear systems related to the stability of orbits of the then known seven planets.

Jacobi did not formulate the problem in the form as described above, since matrix notation was unknown at that time. Actually, he formulated the problem in terms of systems of equations, written out elementwise. His paper contains computational elements that are still in use, like the plane rotations for making a system more diagonally dominant, and the (Gauss)-Jacobi iteration method for diagonally dominant systems. These elements were combined in his approach for the computation of eigenvalues, together with a clever way of setting up a correction equation for eigenvalue and eigenvector approximations.

Later on, other methods became popular, such as the Power iteration method, and, when tools became more sophisticated, eigenproblems other than the standard $Ax = \lambda x$ problem required solution: for instance, the generalized eigenproblems $Ax = \lambda Bx$, higher order polynomial eigenproblems

$$(A_0 + \lambda A_1 + ... + \lambda^n A_n)\, x = 0,$$

and, of course, the singular value decomposition.

Through pioneering work of such people as Arnoldi, Francis, Givens, Householder, Kublanovskaya, Lanczos, Ostrowski, Rutishauser, Wilkinson, and many others, a sophisticated toolbox of algorithms, together with analysis, became available. This work was continued by Kahan, Paige, Parlett, Stewart, to mention only a handful of leaders in a lively and productive research area.

Eigenvalue problems are commonly classified under linear algebra, although they are nonlinear by themselves. Eigenvalues and eigenvectors reveal important properties of the linear operator A and henceforth also of the operator from which A is derived. The nonlinearity makes them special when it comes to numerical computations. The only operations that we can do with a computer are linear and from this one might already conclude that eigenvalue computations are more complicated than, for instance, solving linear systems.

This book is devoted to numerical algorithms for eigenvalue problems. The determination of eigenvalues of A is equivalent with determining the roots of a polynomial, and for that reason there is no finite algorithm that leads to these eigenvalues, for general matrices of order 4 or larger. This explains that all algorithms for eigenvalue computation are essentially iterative. However, there are levels of iterativeness. There are methods that converge very fast, that is quadratical convergence for unsymmetric matrices and cubic convergence for symmetric matrices. With several clever enhancements, these methods are in practice so fast that they lead to all eigenvalues of an arbitrary matrix in a number of floating point operations that is equivalent with the costs for a few matrix matrix multiplications. These methods are clever mixes of the QR method and Divide and Conquer techniques. Since these methods are so fast, they are sometimes referred to as direct techniques in order to distinguish them from the techniques for larger matrices, to be described in subsequent chapters of this book.

Many modern applications lead to eigenvalue problems for matrices of which the order easily exceeds the order of say 10^6, for instance, the study of instabilities in Plasma Physics, or dominant modes in ocean flows, lead to such problems. Problems in Quamtum Chromo Dynamics and Chemistry lead to even much bigger matrices. It is obvious that this market cannot be served with the robust techniques for the smaller order matrices. For these large order matrices, a new level of iteration is used in which the matrix is first reduced to a matrix of much smaller order. This reduced order matrix is then handled by the fast techniques for smaller matrices. The reduced matrix itself is expanded in an iterative way with the aim that some of the eigenvalues of the reduced matrix become good approximations for some of the wanted eigenvalues of the given large matrix. Because these large matrices are often the result of some discretization model, only a part of their eigenvalues, say 10% or less, bears significance for the underlying physical problem. If all eigenvalues and eigenvectors were needed, then these iterative large matrix algorithms also would lead to huge computational times. Methods that fall in the category of reduction algorithms are the Power Method, Simultaneous Iteration, the methods of Lanczos, Arnoldi, and Davidson, and the Jacobi-Davidson methods.

The current state of the art is that most eigenvalue problems of modest order, say a few thousands at most, can be solved by reliable software. The basis for this software was laid in the early 1960s with the publication of Wilkinsons's masterpiece 'The Algebraic Eigenvalue Problem' (WILKINSON [1965]). The mother versions of the actual software were published some ten years later in the famous book edited by WILKINSON and REINSCH [1971], following their earlier publication in Numerische Mathematik. These mother versions were coded in Algol60, a computer language that was popular (in particular in Europe) for numerical computations. Shortly after its publication, this led to the software collection EISPACK, a bunch of Fortran subroutines for standard and generalized eigenproblems. In modern software for eigenvalue problems we still find the traces of the Wilkinson-Reinsch collection.

Why is there a need for still other techniques? Well, the overall costs of these 'direct' algorithms is proportional to n^3, which makes these methods expensive, even on fast supercomputers, when n exceeds a few thousands. Also, the memory requirements are in the order of n^2, which is no surprise for dense matrices, but may be not so attractive for larger sparse matrices.

Because the eigenvalue problem for matrices of modest order may be considered as solved, most of the attention in this book will go to the iterative reduction techniques and our aim is to give the reader a feeling of the state of the art in this field as well as a good idea of the scope of applicability of each of the methods. The main problem with all reduction techniques is that in the reduction we discard information of the large given system and then we have to decide in the process whether the discarded information had still some significance for the approximations that we compute from the reduced system. This leads to such phenomena as *misconvergence*, where, for instance, the largest eigenvalue of the reduced matrix is not close to the (wanted) largest eigenvalue of the large matrix. We will try to give the reader an idea of when such situations can happen and what can be done to improve our trust in the approximation quality of computed eigenvalues (and eigenvectors).

2 The history of eigenvalue algorithms

Major part of this section has been taken from Golub and van der Vorst [2001] and van der Vorst and Golub [1997].

Early computational techniques, other than Jacobi's famous but slowly converging diagonalization method and the unsatisfactory Power method with its many restrictions, attempted to exploit the fact that every matrix satisfies its characteristic equation. To this end, Krylov suggested in 1931, using the vectors x, Ax, A^2x, ..., generated by the Power method, to determine the coefficients of the characteristic equation. This was not successful, because, as we have learned from the analysis in Wilkinson [1965], the roots of a polynomial may vary widely with only tiny perturbations to the coefficients of the polynomial. Even rounding the exact coefficients in floating point arithmetic may destroy much accuracy in many of the roots. Although Krylov's method failed, his name is still attached to the subspace generated by the Power method.

There is yet another reason for the failure of Krylov's method in finite precision arithmetic: the vectors generated by the Power method tend to converge in the direction of the eigenvectors associated with the dominating eigenvalues. Hence, the computed vectors for the subspace necessarily yield a very ill-conditioned basis. Checking mutual dependence of this basis, as is required in order to construct the characteristic polynomial, is an almost impossible task.

In 1952, Lanczos suggested building a basis for the Krylov subspace

$$\mathcal{K}^m(A;x) := \mathrm{Span}\{x, Ax, A^2x, \ldots, A^{m-1}x\}$$

in a more stable way by orthogonalizing the vectors as they are constructed. The idea was to immediately orthogonalize the new vector for the Krylov subspace with respect to the already existing orthonormal basis. The new vector for expansion is created by applying A to the latest orthogonal basis vector. For symmetric matrices A, this leads to a three-term recurrence relation between the basis vectors v_j and in exact arithmetic this can be formulated as

$$AV_j = V_j T_j + \gamma_j v_{j+1} e_{j+1}^T.$$

Obviously, this recursion must terminate for some $j \leq n$, in which case V_j forms the basis for an invariant subspace of A, and the eigenvalues of the tridiagonal j by j matrix T_j are the eigenvalues of A with respect to this invariant subspace. This algorithm is equivalent to the well-known algorithm of the Dutch mathematician Stieltjes (1856-1895) for generating orthogonal polynomials by a three-term recurrence relationship. Lanczos also proposed a reduction process for unsymmetric matrices A, the so-called two-sided Lanczos process. In this process two sets of basis vectors are constructed, one for the Krylov subspace with A and one for a Krylov subspace with A^T. By requiring bi-orthogonality of the two sets, the two bases can be used for reduction of A to tridiagonal form. This form has suffered from many drawbacks. Not only are the reduction matrices non-orthogonal, a suspect property, but the algorithm also suffers from various break-down conditions. (The basic problem lies in the fact that the measure generated by the initial vectors is not non-negative). The symmetric variant did not become popular in the 1950s, since it was soon recognized that rounding errors could spoil the process dramatically. WILKINSON [1965] showed that the Lanczos algorithm is highly (forward) unstable and there seemed to be no way of stabilizing the process other than re-orthogonalizing the generated vectors. He showed that this process is comparable to the methods of Householder or Givens (proposed in the late fifties), but the latter are more economical. He then concluded: "*it is difficult to think of any reason why we should use Lanczos' method in preference to Householder's.*" This illustrates that the Lanczos method was commonly viewed as a direct reduction method at that time, and from that point of view Wilkinson's remarks were quite correct.

At about the same time as Lanczos, Arnoldi (1951) gave a reduction algorithm for unsymmetric matrices. This was basically the same algorithm as Lanczos' algorithm for symmetric matrices, with the difference that each new basis vector had to be orthogonalized with respect to all previous basis vectors. In this way A is reduced by an orthogonal similarity transformation to upper Hessenberg form. Arnoldi's method suffers far less from numerical instability, depending on how well the orthogonalization process is carried out. But the method is more expensive than the Householder reduction, making it less attractive, as a direct method, for large (dense) matrices.

A very important notion was the recognition that matrices could be reduced, by orthogonal transformations, in a finite number of steps, to some special reduced form that lends itself more efficiently to further computations. In particular, a symmetric matrix can be reduced to tridiagonal form by Jacobi-rotations, provided that these rotations are restricted to annihilate entries of A outside its tridiagonal part. This was suggested by Givens in 1954, and in this connection the Jacobi-rotations are also called Givens rotations. Almost simultaneously, Householder, in 1958, discovered that complete columns of A could be reduced to zero, outside the tridiagonal part, by the more efficient Householder reflections. These are well-chosen orthogonal rank-one updates of the form

$$(I - \frac{2}{v^T v} v v^T)\ ;$$

they are discussed in more detail in TURNBULL and AITKEN [1961]. The Householder method has become the method of choice for the reduction of matrices to tridiagonal form on serial computers. Thus for eigenproblems, a symmetric matrix can be reduced by a finite number of orthogonal similarity transformations to tridiagonal form and unsymmetric matrices can be transformed to upper Hessenberg form.

By 1960, the eigenvalue problem for a symmetric tridiagonal matrix was solved by using the Sturm sequence property for successive subdeterminants. The corresponding eigenvectors were computed by inverse iteration. The whole process is described in Givens' papers. A complete and thorough analysis for the Givens and Householder reductions and for the use of the Sturm sequences, is given in Wilkinson's book, which was the numerical linear algebra bible for a long time.

As we will see in Section 10, the QR method is, for determining the complete set of eigenvalues and eigenvectors, a superior technique. At the time that Wilkinson's book appeared, the blossoming of the QR method had just begun. Wilkinson devoted much attention to this method, but not as the method of choice for symmetric problems. We quote from PARLETT [1980, p. 172]: "*Yet it was not invented until 1958-1959 and was not appreciated until the mid-1960s. The key idea came from Rutishauser with his construction of a related algorithm called LR in 1958.*" Whereas Wilkinson's book was the reference for eigenvalue problems in the period 1960-1980, after 1980, Parlett's book *The Symmetric Eigenvalue Problem* became the main source, at least for symmetric problems. Comparison of the two books clearly shows the progress made in this field.

The use of the QR (a mathematical equivalent is the QL algorithm) algorithm began with the work of FRANCIS [1961], who recognized in 1961-1962 that a QR iteration maps a Hessenberg matrix to a Hessenberg matrix again, and this makes the process economical and also adds to stability since the zero elements need not be computed. Furthermore, Francis cleverly used implicit origin shifts, and these can be carried out very economically for Hessenberg matrices. Kublanovskaja, in 1961, discovered independently the same process, but did not employ the invariance of the Hessenberg form. She deeply understood the mathematical aspects of the algorithm but was less concerned with the important computational details. The inclusion of Wilkinson shifts eventually makes the process very efficient, and for these shifts it can be proved that, for symmetric matrices, the process does not fail. The order of convergence for symmetric matrices is cubic (see, for instance, HOFFMANN and PARLETT [1978]), while for unsymmetric matrices it is quadratic (see GOLUB and VAN LOAN [1996] or DEMMEL [1997]). These results rest on work of Ostrowski carried out in connection with the shift-and-inverse Power Method (the RQI method, see that section).

For a treatment of modern implementations of the QR method see GOLUB and VAN LOAN [1996, Chapter 7] or DEMMEL [1997, Section 4.4.5]. These implementations incorporate techniques developed in the 1990s, such as (double) implicit shifts. This implicit shift technique leads to a rank-one perturbation of the Hessenberg structure, and this perturbation can be removed in turn by a technique that is known as *chasing the bulge*: the perturbation (bulge) is chased down (and out of) the Hessenberg matrix with (multiple) Givens rotations. These chasing techniques were analysed by WATKINS and ELSNER [1991]. An important and complete overview of the practical QR algorithm can be found in WATKINS [1993].

With respect to the relation between the QR and the Lanczos algorithms, we note the following. Lanczos' method focusses on one particular starting vector; QR starts with a full orthogonal basis and keeps it orthogonal through the Power iterations; the inclusion of shifts does not destroy the structure of the Hessenberg matrix. With Lanczos' method, a shift only makes sense in damping unwanted parts of the spectrum, but one cannot vary the shift during the process.

By 1970 the standard numerical eigenproblem, for dense matrices of not too large order, could be regarded as essentially solved and research shifted to larger problems and other eigenproblems. The next important problem to consider was the generalized eigenproblem $Ax - \lambda Bx = 0$. An obvious approach is to transform this to a standard eigenproblem by inverting either A or B, or to work with more complicated transformations, such as the

Cayley Transform: $(A - \sigma B)^{-1}(A - \tau B)$. These approaches share the disadvantage that matrices A and B are not treated in the same way, which is most obvious from the simple transformation $B^{-1}A$. This leads to problems if B is singular or ill-conditioned, but as Stewart (1980) has pointed out, this does not necessarily mean that the given eigenproblem is ill-conditioned because if A is well-conditioned then the pencil $B - \mu A$ may still be well behaved (small perturbations in A and B lead to small perturbations in μ).

In Stewart's analysis the matrices are treated symmetrically; in particular he suggests considering the pencil $\sigma A - \tau B$, and regarding multiples of (σ, τ), for which the determinant $|\sigma A - \tau B|$ vanishes, as generalized eigenpairs. See STEWART and SUN [1990, Chapter VI] for more details. This approach, of course, requires a different reduction and this is accomplished by the Generalized Schur Decomposition, proposed by Moler and Stewart in 1973. This says that for arbitrary square A and B there exist unitary Q and Z such that $Q^*AZ = T$ and $Q^*BZ = S$ are upper triangular. For real matrices, the arithmetic can be kept real, but then the reduced matrices are quasi-triangular (that is 2 by 2 non-zero blocks along the diagonal may occur).

MOLER and STEWART [1973] also proposed a stable algorithm to accomplish the reduction to (quasi) triangular form and this is known as the *QZ algorithm*. This algorithm will be discussed in some detail in Chapter 12. Major modern software packages include software for the QZ algorithm. For perturbation analysis, we refer to STEWART and SUN [1990].

With larger problems coming into the picture, it soon became clear that the matrix transforming techniques could not solve these problems with reasonable computing resources, and, as an alternative, iterative methods were investigated. The Power method was not robust enough, although the method is still hidden in state-of-the-art methods, like the powerful (literally) QR method, see, for instance, GOLUB and VAN LOAN [1996] or PARLETT [1980]. LANCZOS [1950] and ARNOLDI [1951] started the research on modern iteration methods in the early 1950s.

After a period of little interest in these methods, mainly because of poor understanding of their numerical properties, PAIGE [1972] and PAIGE [1976] showed the potential of the Lanczos method. This marked the start of an entire new area of research. About the same time DAVIDSON [1975], a chemist, designed a popular method that became widely known as Davidson's method. Among numerical analysts the method did not receive much attention, partly because it appeared to be applicable only to diagonal dominant problems, and partly because its effects were poorly understood from a numerical point of view.

The Lanczos method, for symmetric matrices, was well understood by the mid 1980s, see for instance Parlett [1980], but the methods for unsymmetric systems have begun to blossom by that time.

It should be noted that in the modern iterative *subspace* methods, like Arnoldi's, Lanczos', and Davidson's method, the given large problem is reduced to a much smaller problem. This smaller problem can then be solved by the by now standard techniques for dense matrices, available through *netlib*, in *Matlab*, *LAPACK*, and in major numerical software libraries.

In the past ten years we have seen many new developments in this exciting area. In our opinion, the major steps, from an algorithmic point of view, are:

- Attempts to make the Two-sided Lanczos process more robust, by introducing a so-called Look-ahead strategy (Parlett, Taylor and Liu [1985], Freund and Nachtigal [1991], Brezinski and Redivo-Zaglia [1994], Gutknecht [1992]).

- The idea of an implicit restart technique for the Arnoldi process, which helps to keep memory requirements reasonable, and which makes the Arnoldi process an attractive algorithm for eigencomputations (Sorensen [1992]).

- Further improvements on the Davidson method, culminating in the Jacobi-Davidson algorithm (Sleijpen and van der Vorst [1996]). In this algorithm a major deficiency in the original Davidson method has been removed. The Jacobi-Davidson method can be implemented as an accelerated inner-outer iteration scheme.

Other important developments and improvements were (the list is still personally colored and incomplete):

- The concept and use of the ε-pseudospectra of a matrix A (Trefethen [1992]) as a means to make sensitivities in the spectrum of a nonnormal matrix clearly visible. The ε-pseudospectra approach does not make classical perturbation theory superfluous. It helps to detect situations that need further analysis, and it does so in a way that is easily understandable for non-expert users. We will discuss the use of this tool briefly.

- The convergence behavior of Ritz values, for symmetric matrices, has been further analyzed and is now quite well understood. The so-called misconvergence phenomenon, a Ritz value that lingers near some eigenvalue before it converges towards some other eigenvalue (van der Sluis and van der Vorst [1987a], Parlett [1990]), and various other local effects in the convergence behavior have been analysed and are fairly well understood (van der Sluis and van der Vorst [1987a]). The notion that rounding errors can lead to multiple (spurious) eigenvalues, has been translated into practical strategies, not only for the Lanczos method, but also for the Two-sided Lanczos process (Cullum and Willoughby [1986]).

- Strategies have been proposed to improve the convergence behavior of the various subspace methods. These strategies all amount to filtering undesired eigenvector components from the starting vector: Chebychev polynomial preconditioning (Saad [1992, Chapter VII]), rational shift-and-invert techniques (Ruhe [1994a], Ruhe [1994b]), or to increase desired eigenvector components (inexact shift-invert Preconditioning for Eigenproblems), see Morgan and Scott [1993] and Meerbergen [1996], and Saad [1992, Chapter VII]. The general idea goes back to Flanders and Shortley [1950]; Lanczos [1956] also suggested polynomial filtering techniques.

- Generalizations of the Davidson method, making the method also suitable for unsymmetric matrices. These methods come down either to the incorporation of more general preconditioning, instead of the originally proposed diagonal preconditioning (Morgan and Scott [1986], Morgan [1992], Crouzeix, Philippe and Sadkane [1994]), or to second order corrections for the current eigenvector approximations (Olsen, Jørgensen and Simons [1990]. The latter variant is strongly related to the Jacobi-Davidson type of algorithms.

- Eigenproblems can be regarded as nonlinear problems, which means that one can employ homotopy methods for the efficient computation of eigenvalues, or tracking these eigenvalues as function of a parameter in the underlying model (Lui and Golub [1995], Lui, Keller and Kwok [1997]). We will sketch this approach and point at some relations with other iterative methods.

- Parallelism (Cuppen's Divide and Conquer, see Cuppen [1981], Dongarra and Sorensen [1987], and restructuring of iterative algorithms in an attempt to combine innerproducts, see Demmel, Heath and van der Vorst [1993], Bai, Hu and Reichel [1991], De Sturler and van der Vorst [1995]).

- Subspace methods for interior eigenvalues (MORGAN and SCOTT [1986], RUHE [1994b]), rational Lanczos (RUHE [1994a], RUHE [1994b]), harmonic Ritz values (FREUND [1992], MORGAN [1991], PAIGE, PARLETT and VAN DER VORST [1995]).

Interior eigenvalues are always difficult to compute with subspace methods, if one wants to avoid expensive shift-and-invert operations. Approximate shift-and-invert operations have been suggested for the Jacobi-Davidson method in SLEIJPEN and VAN DER VORST [1996], for the Lanczos method in MORGAN and SCOTT [1993], and for the Arnoldi method in MEERBERGEN [1996]. The notion of harmonic Ritz values offers a helpful tool for restart purposes, since they identify the best approximations with respect to interior eigenvalues.

3 Layout of this article

In this section we will give a brief overview of what the reader may expect to see in this book.

In Chapter II we review some important basic properties that are required throughout the remainder of the text. Usually, a given matrix is transformed to some more convenient form and these canonical forms are given together with conditions for their existence. Furthermore, we give some of the most important results for perturbations of eigenvalues and eigenvectors.

For historical reasons, we start with a description of Jacobi's algorithm in Chapter III. Not only is this algorithm a very elegant one, but parts of it still play a role in other algorithms as well. Moreover, the method has attractive properties with respect to parallelism and numerical accuracy.

In Chapter IV the so-called 'direct' methods are presented. These methods, like most of the other methods, with the exception of Jacobi's method, use variations on the Power Method. Therefore, we start that chapter with this (often rather slowly converging) iterative method, see section 8. One of the most elegant algorithms in this field is the QR method. It is a subtle composition of by itself simple ingredients, as we will explain in section 10. The generalization of QR to the generalized eigenproblem $Ax = \lambda Bx$ is known as the QZ method. This method requires a few additionel tricks to make it work efficiently and accurately, see section 12.

Many truly iterative methods are based on the approach of projecting the given standard eigenproblem onto a so-called Krylov subspace. Because of the importance of the Krylov subspace and because of its relevance for most of the iterative methods to come, we have added a separate chapter for a discussion on this matter (Chapter V). In that chapter, we give the most relevant properties of the Krylov subspace. We also briefly introduce the eigenvalue and eigenvector approximations computed from these subspaces: the so-called *Ritz values* and *Ritz vectors*. Furthermore, we give an overview of some other aspects, such as the *Harmonic Ritz values* and *Harmonic Ritz vectors*, the *Refined eigenvector approximations*, and Block variants.

For the remaining chapters, we have to distinguish more strongly between Hermitian and non-Hermitian matrices, because in contrast to a Hermitian matrix a non-Hermitian one does in general not have an orthogonal set of eigenvectors, or in other words, a non-Hermitian matrix A can in general not be transformed to diagonal form $D = Q^*AQ$ by an orthogonal matrix Q. This has consequences for the selection of algorithms and for the stability of the computed results, even more than for direct methods. Conceptually, one can apply the QR method for general matrices, but the iterative methods are often restricted to matrices with special properties. Especially Hermitianicity plays an important role.

For large Hermitian matrices the standard method of choice is the Lanczos method. Chapter VI offers a detailed discussion on various algorithmical and theoretical aspects of this method. The Lanczos method may be regarded as a special case of the Arnoldi method: it generates an orthonormal basis for the Krylov subspace with short recurrences. Indeed, if $A^* = A$, then from

$$AV_m = V_{m+1}H_{m,m+1}$$

it follows that $V_m^* AV_m = H_{m,m}$ is Hermitian, and hence $H_{m,m}$ is a Hermitian tridiagonal matrix. In order to realize fast convergence for the Ritz values towards eigenvalues close to a target σ, one often applies the Lanczos method with the operator $(A - \sigma I)^{-1}$.

For general unsymmetric matrices, the Krylov subspace concept leads to the Arnoldi method. In Chapter VII, we will discuss important properties of this method, for the computation of selected eigenvalues and eigenvectors. In particular, we will highlight the elegant *Implicit Restarted Arnoldi Method* as a means to achieve convergence to selected eigenvalues with subspaces of modest dimensions.

The Arnoldi methods require the construction and storage of a full basis for the Krylov subspace. This becomes increasingly expensive with increasing dimension of the subspace. For Hermitian A there is a simple three-term recurrence that generates a basis for the subspace, and if one is interested in the eigenvalues only, then only the last three basis vectors have to be stored.

A variant of this *Lanczos* process (see Chapter VI) for non-Hermitian matrices is known as the *Two-sided Lanczos Method* and this is the subject of Chapter VIII. For this variant one has to give up the orthogonality of the basis vectors. Hence economy in memory and computational overhead comes at the risk of instability. The original Two-sided Lanczos method suffers also from (near) break-downs. Several variants have been suggested that have been proven to be efficient and useful for relevant problems.

The Arnoldi and Lanczos methods are most effective when applied with shift and invert, that is systems of the form $(A - \sigma I)x = b$ have to be solved efficiently and accurately. If neither of this is practical, but if some reasonable approximation for $x = (A - \sigma I)^{-1}$ is cheaply available, then the Jacobi-Davidson methods can be useful. These methods work with orthogonal transformations, just as the Arnoldi methods, but in the construction of an orthonormal basis for the involved subspace there is no attempt to create a special structure for the reduced matrix. Therefore, the computational overhead is larger than for Arnoldi. All the possible gain has to come from an effective and cheap approximation for the shift-and-invert step. The Jacobi-Davidson methods are described in Chapter IX. That chapter includes also a discussion on the original Davidson method.

Iterative methods for the generalized eigenproblem are slightly more complicated. They are the subject of discussion in Chapter X. An obvious approach is to transform the generalized eigenproblem $Ax = \lambda Bx$ to a related convenient standard eigenproblem $Cy = \mu y$. We will review several possiblities. Arnoldi's method has a variant that is special geared for the generalized eigenproblem, the *Rational Krylov Sequence* approach. Furthermore, we will pay much attention to applications of the Jacobi-Davidson approach. This approach helps to avoid transformation and permits to work with Schur vectors. Both aspects help to improve numerical stability.

In Chapter XI we consider some unusual eigenproblems. More in particular, we pay attention to polynomial eigenproblems

$$(A_0 + \lambda A_1 + ..\lambda^n A_n)\, x = 0,$$

with the second order case as an important subclass. Furthermore, we will pay some attention to constrained eigenproblems and their numerical solution.

Finally, we have gathered some related issues in Chapter XII. In particular, we give a brief overview of the concept of *pseudospectra*, an issue that is highly related with stability of the eigenproblem. Also, homotopy methods receive attention as a means to compute some eigenvalues for 'difficult', or perturbed matrices, starting with available knowledge for a given matrix. We shall briefly examine such techniques. The chapter is concluded with some remarks on implementation aspects of iterative methods for the eigenvalue problem.

Acknowledgements For the preparation of the text of this article, I received valuable help of Jan Brandts, Jos van Dorsselaer, and Gerard Sleijpen. They pointed at errors and poor formulations in earlier drafts and they suggested numerous improvements. Jos van Dorsselaer provided the example with the field of values for Arnoldi's method (Chapter VII).

CHAPTER II

Basic Theory

4 Canonical forms

The standard approach for the numerical solution of the eigenproblem is to reduce the operators involved to some simpler form, that reveals the eigenvalues and eigenvectors directly, for instance, diagonal form. The idea is that the transformation is carried out, as much as possible, with unitary operators, in order to reduce the effects of perturbations.

The easiest situation is the symmetric case: for a real symmetric matrix, there exists an orthogonal matrix Q, so that $Q^T AQ = D$, where D is a diagonal matrix. The diagonal elements of D are the eigenvalues of A, the columns of Q are the corresponding eigenvectors of A.

Unsymmetric matrices do in general not have an orthonormal set of eigenvectors, and may not have a complete set of eigenvectors, but they can be transformed unitarily to Schur form:

$$Q^* AQ = R,$$

in which R is upper triangular. In fact, the symmetric case is a special case of this Schur decomposition, since a symmetric triangular matrix is clearly diagonal. Apart from the ordering of the eigenvalues along the diagonal of R and the sign of each column of Q, the matrix Q is unique. VAN DOOREN [1982] has described an algorithm for the permutation of Q, so that the eigenvalues appear in prescribed order along the diagonal of R.

If the matrix A is complex, then the matrices Q and R may be complex as well. However, they may be complex even when A is real unsymmetric. It may then be advantageous to work in real arithmetic. This can be realized because of the existence of the *real Schur decomposition*. If $A \in \mathbb{R}^{n\times n}$ then it can be transformed with an orthonormal $Q \in \mathbb{R}^{n\times n}$ as

$$Q^T AQ = \widetilde{R},$$

with

$$\widetilde{R} = \begin{bmatrix} \widetilde{R}_{1,1} & \widetilde{R}_{1,2} & \cdots & \widetilde{R}_{1,k} \\ 0 & \widetilde{R}_{2,2} & \cdots & \widetilde{R}_{2,k} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \widetilde{R}_{k,k} \end{bmatrix} \in \mathbb{R}^{n \times n}.$$

Each $\widetilde{R}_{i,i}$ is either 1 by 1 or a 2 by 2 (real) matrix having complex conjugate eigenvalues. For a proof of this see GOLUB and VAN LOAN [1996, Chapter 7.4.1]. This form of $\widetilde{R}$ is referred to as an upper *quasi-triangular matrix*.

If all eigenvalues are distinct then there exists a nonsingular matrix X (in general not orthogonal) that transforms A to diagonal form:

$$X^{-1}AX = D.$$

A general matrix can be transformed to Jordan form with a nonsingular X:

$$X^{-1}AX = \mathrm{diag}(J_1, J_2, \ldots, J_k),$$

where

$$J_i = \begin{bmatrix} \lambda_i & 1 & 0 & \cdots & 0 \\ 0 & \lambda_i & \ddots & & \vdots \\ & \ddots & \ddots & \ddots & \\ \vdots & & \ddots & \ddots & 1 \\ 0 & \cdots & & 0 & \lambda_i \end{bmatrix}.$$

If there is a J_i with dimension greater than 1 then the matrix A is defective. In this case A does not have a complete set of independent eigenvectors. In numerical computations one may argue that small perturbations lead to different eigenvalues and hence that it will be unlikely that A has a true Jordan form in actual computation. However, if A is close to a matrix with a non-trivial Jordan block, then this is reflected by a (severely) ill-conditioned eigenvector matrix X.

The numerical computation of the Jordan form is highly unstable, since a small perturbation suffices to obtain a matrix with different eigenvalues (and possibly a complete eigenvector system). Small angles between (some of) the eigenvectors reveal that A is close to a matrix that is similar to a non-diagonal Jordan form. For a discussion on how to compute elementary Jordan blocks (with the help of the Singular Value Decomposition), see the paper by GOLUB and WILKINSON [1976].

If possible, we prefer to work with well-conditioned matrices, and with unitary ones in particular, and for this reason we prefer to work with the (real) Schur decomposition for (real) unsymmetric matrices:

$$Q^T AQ = \widetilde{R},$$

where $\widetilde{R}$ has upper quasi-triangular form. The 1 by 1 blocks along the diagonal of $\widetilde{R}$ represent the real eigenvalues of A, the 2 by 2 blocks represent the complex conjugate eigenvalues of the real matrix A. An eigenvector x of A corresponding to λ follows from the solution of

$$\widetilde{R}y = \lambda y, \quad \text{with} \quad x = Qy,$$

and this formulation makes it much easier to detect the problems associated with the numerical computation of eigenvectors, such as near-defectiveness.

Just as the Jordan canonical form describes the eigenstructure of a matrix, the Kronecker form does this for matrix pencil $A - \lambda B$, even for rectangular A and B. For details on this we refer to WILKINSON [1978], WILKINSON [1979], VAN DOOREN [1979], and KÅGSTRÖM [1985]. In KÅGSTRÖM [1986] is also software for the computation of the Kronecker discussed. Wilkinson, in his discussion on the progress made in the period 1976-1986 in eigenvalue computations (WILKINSON [1987]), noted that the Jordan canonical and Kronecker canonical forms were largely regarded as irrelevant by numerical analysts because of their ill-posedness. He even stated: "*Many felt that I should have ignored the Jordan canonical form in the Algebraic Eigenvalue Problem (*WILKINSON *[1965]) and I had misgivings about including a discussion of it*". Since the 1970s, this has changed, and contributions have been made by many, including Demmel, Beelen, Van Dooren, Chaitin-Chatelin, Edelman, Kågström, Nichols, Kautsky, Golub, and Wilkinson.

Although serious attempts have been undertaken for the computation of the Kronecker canonical form, for instance by Kågström and Van Dooren, this still needs further research. Also the computation of invariant subspaces of highly nonnormal matrices is still in its infancy, notwithstanding useful contributions by, for instance, CHAITIN-CHATELIN [1994], BRACONNIER, CHATELIN and FRAYSSÉ [1994], and LEE [1995]. For recent references, see BAI *et al* [2000]. Van Dooren described, in papers published in 1981, how the Kronecker form can be used in system control problems (input-output systems), see VAN DOOREN [1981a], VAN DOOREN [1981b].

For matrix pairs $A, B \in \mathbb{C}^{n\times n}$ there exist a *generalized Schur decomposition*:

$$Q^* AZ = R \quad \text{and} \quad Q^* BZ = S,$$

with unitary $Q, Z \in \mathbb{C}^{n\times n}$ and the $R, S \in \mathbb{C}^{n\times n}$ are upper triangular matrices (MOLER and STEWART [1973]). If $A, B \in \mathbb{R}^{n\times n}$. then all arithmetic can be kept real because of the existence of a *generalized real Schur decomposition.* In that case, there exist orthogonal matrices Q and Z such that

$$Q^T AZ = \widetilde{R} \quad \text{and} \quad Q^T BZ = \widetilde{S},$$

with $\widetilde{R}, \widetilde{S} \in \mathbb{R}^{n\times n}$ upper quasi-triangular matrices (STEWART [1972]).

The famous Singular Value Decomposition (SVD) is closely related to eigendecompositions. Let A be a real m by n matrix, then there exists an orthogonal m by m matrix U and an orthogonal n by n matrix V, such that

$$U^T AV = \operatorname{diag}(\sigma_1, \ldots, \sigma_p), \tag{4.1}$$

with $p = min\{m, n\}$, and $\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_p \geq 0$. The values σ_i are the *singular values*, , the columns v_i of V are the *right singular vectors* and the columns u_i of U are the *left singular vectors* of A. The number of nonzero singular values is equal to the rank of the matrix A. In finite precision arithmetic it rarely happens that singular values are equal to zero, even if they should have been in exact arithmetic. Therefore, one also uses the notation of numerical rank of a matrix. If for some $\epsilon > 0$, the singular values can be ordered as

$$\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_t \geq \epsilon > \sigma_{t+1} \geq \ldots \geq \sigma_p \geq 0,$$

then we say that A has *numerical rank* t (with respect to ϵ).

The parallel with eigenpairs of a square matrix A is made visible through writing out (4.1) for the i-th column of AV, and for the i-th column of $A^T U$:

$$\begin{aligned} Av_i &= \sigma_i u_i \\ A^T u_i &= \sigma_i v_i. \end{aligned} \tag{4.2}$$

Together, these equations can be expressed as a true eigenvalue problem for an auxiliary matrix:

$$\begin{bmatrix} 0 & A \\ A^T & 0 \end{bmatrix} \begin{bmatrix} u_i \\ v_i \end{bmatrix} = \sigma_i \begin{bmatrix} u_i \\ v_i \end{bmatrix}. \tag{4.3}$$

Other relations, that are of interest in order to understand properties of the singular vectors and singular vectors are

$$A^T Av_i = \sigma v_i, \tag{4.4}$$

$$AA^T u_i = \sigma u_i. \tag{4.5}$$

Note that these equations are not always very adequate for the computation of accurate singular vectors and singular values, because of the matrix products A^TA and AA^T, which may lead to significantly larger perturbations.

The SVD can also be defined for a complex matrix $A \in \mathbb{C}^{m\times n}$. The only difference with the real case is that unitary matrices U and V are involved in the reduction to diagonal form:

$$U^H AV = \mathrm{diag}(\sigma_1, \ldots, \sigma_p).$$

The SVD plays an important role in numerical linear algebra, for instance in the solution of underdetermined or overdetermined linear systems. For a good overview of the numerical properties of the SVD, see GOLUB and VAN LOAN [1996].

5 Perturbation results

Perturbation theorems play a very essential role in computational processes for eigenproblems. As we have remarked, eigensolvers are essentially iterative processes and many of them rest on the principle of reducing a matrix to a special/simpler form, either diagonal or upper triangular. One has then to decide when a matrix is sufficiently close to the appropriate ultimate form. It is important to know what the approximation errors imply about the desired eigeninformation. A modern treatment of perturbation theory for a variety of eigenproblems is given in the book by STEWART and SUN [1990]. We will restrict ourselves here to what we regard as some of the most relevant results in this area.

Many theoretical results rest on the famous Gerschgorin Disc Theorem, which states that the eigenvalues are in discs around the diagonal elements:

Theorem 1 *Let $A = D+F$, where $D = diag(A) = diag(d_1, d_2, \ldots, d_n)$ and F has zero diagonal entries.*
Then,

$$\lambda(A) \subseteq \bigcup_{i=1}^{n} D_i$$

where

$$D_i = \left\{ z \in \mathbb{C} : |z - d_i| \leq \sum_{j=1}^{n} |f_{ij}| \right\}.$$

This theorem first appeared in a classic paper by GERSCHGORIN [1931]. A very useful refinement shows that if a group of s discs D_i is isolated from the others, then this group contains precisely s eigenvalues. In particular, if one disc is isolated then this disc contains one eigenvalue. This particular case is of great interest, since it can be used for stopping criteria in actual computations.

WILKINSON [1965] discussed the application of Gerschgorin's theorem to various situations. He mentioned the discussion of useful extensions by TAUSSKY [1949]: VARGA [1962] acknowledged Taussky's paper in his work on solving systems of linear equations by iterative methods. An important extension, using block matrices, of the Gerschgorin Theorem was given by FEINGOLD and VARGA [1962].

The eigenvalues depend continuously on the entries of A and if the ith eigenvalue λ_i is distinct, then it is even differentiable. In this case one can carry through a first order perturbation analysis (cf. STEWART and SUN [1990, p. 185]). This leads to the observation that if a matrix A is perturbed by δA, then the perturbation to λ_i is in first order of terms of δA given by

$$\delta\lambda_i = \frac{1}{y_i^* x_i} y_i^* \delta A x_i, \tag{5.1}$$

where x_i, and y_i are right and left eigenvectors respectively, corresponding to λ_i, and y_i^* denotes the complex conjugate of y_i. The factor $\dfrac{1}{y_i^* x_i}$ is referred to as the *condition number* of the ith eigenvalue. The result of BAUER and FIKE [1960], which is actually one of the more well-known refinements of Gerschgorin's theorem, makes this more precise: the eigenvalues $\widetilde{\lambda}_j$ of $A + \delta A$ lie in discs B_i with center λ_i, and radius $n\dfrac{\|\delta A\|_2}{|y_i^* x_i|}$ (for normalized x_i and y_i).

As we will see, many computational methods for larger matrices use the idea of restricting the matrix to some convenient subspace and this leads to low dimensional restricted matrices. The eigenvalues of such relatively small matrices can be conveniently computed by standard techniques, in particular by the QR method. The idea in such subspace methods is to expand the subspace in order to optimize, for instance, some function of the wanted eigenpair. If the matrix A is symmetric, then the eigenvalues can be characterized by the classical Courant-Fischer theorem, also referred to as the minimax theorem.

Before we can state this theorem, we define the *Rayleigh quotient* $\rho(A, x)$, for vectors $x \neq 0$, as

$$\rho(A, x) = \frac{x^* A x}{x^* x}. \tag{5.2}$$

Let $\lambda_n \leq \lambda_{n-1} \leq \ldots \leq \lambda_1$ denote the (real) eigenvalues of the symmetric matrix A and let x_i denote the corresponding eigenvectors. Then it is easy to prove that

$$\lambda_1 = \max_{0 \neq x \in \mathbb{R}^n} \rho(A, x)$$

and

$$\lambda_n = \min_{0 \neq x \in \mathbb{R}^n} \rho(A, x).$$

The max and min are attained for x_1, x_n, respectively.

With $\mathbb{R}^n = < x_1 > \oplus \mathcal{V}$, we exclude x_1 from $\mathbb{R}^n$ to define $\mathcal{V}$, and then we can formulate a similar characterization for λ_2, as for λ_1, with respect to $\mathcal{V}$. The extension to other eigenvalues is obvious.

Note that in the above sketched approach, it is necessary to have, for instance, x_1, before we can start the maximization of $\rho(A, x)$ in order to determine λ_2 and x_2. This drawback is avoided in the formulation of the minimax theorem and this will prove to be of great help for the derivation of eigenvalue approximations from generated subspaces.

Theorem 2 *(Courant-Fischer Minimax Theorem) Let $\mathcal{V}^j$ denote a linear subspace of dimension j of $\mathbb{R}^n$, and let $A \in \mathbb{R}^{n \times n}$ be symmetric. Then for $j = 1, \ldots, n$, we have that*

$$\lambda_j = \max_{\mathcal{V}^j} \min_{0 \neq x \in \mathcal{V}^j} \rho(A, x),$$

and also

$$\lambda_j = \min_{\mathcal{V}^{n-j+1}} \max_{0 \neq x \in \mathcal{V}^{n-j+1}} \rho(A, x).$$

We have hinted that this theorem can be proven by using the eigenvectors of A. Indeed, if one expands the vector x in terms of the orthonormalized eigenvectors as

$$x = \sum_{j=1}^{n} \gamma_j x_j,$$

then it is rather straightforward to prove the above results. For proofs and more background we refer to WILKINSON [1965, Chapter 2.43].

For symmetric matrices, Weyl in 1912 proved an important consequence for symmetric perturbations δA:

$$|\widetilde{\lambda}_i - \lambda_i| \leq \|\delta A\|_2.$$

In fact, Weyl gave even more detailed results in terms of the eigenvalues of A and of δA. Denote the eigenvalues of δA by $\eta_n \leq \ldots \leq \eta_2 \leq \eta_1$, then the eigenvalues $\widetilde{\lambda}_i$ of $A + \delta A$ satisfy

$$\lambda_i + \eta_n \leq \widetilde{\lambda}_i \leq \lambda_i + \eta_1. \tag{5.3}$$

Another important property that plays a role in iterative (projection type) algorithms, is the interlace property. Let A_r denote the leading r by r minor of A, with eigenvalues $\lambda_j^{(r)}$, then

$$\lambda_{r+1}^{(r+1)} \leq \lambda_r^{(r)} \leq \lambda_r^{(r+1)} \leq \cdots \leq \lambda_2^{(r+1)} \leq \lambda_1^{(r)} \leq \lambda_1^{(r+1)}.$$

An important result, that underlies the powerful *divide and conquer* method (see Chapter 13) , comes from rank-one perturbations. For the formulation of the following result, we will use the convention that $\lambda_{n+1} = -\infty$, $\lambda_0 = \infty$. If $B = A + \tau c c^T$, with $\|c\|_2 = 1$, and real $\tau \geq 0$, then the ith eigenvalue of B is in the interval $[\lambda_{i+1}, \lambda_i]$; for $\tau \leq 0$ it is in $[\lambda_i, \lambda_{i-1}]$. In either case, there exist nonnegative $\alpha_1, \ldots, \alpha_n$ with $\sum_i \alpha_i = 1$, such that $\lambda_i(B) = \lambda_i + \alpha_i \tau$.

Further details and results can be found in most books on the (numerical) eigenproblem; in particular WILKINSON [1965] is a great source. A good overview of these and similar results for invariant subspaces is given in GOLUB and VAN LOAN [1996].

From a result referred to as the Kato-Temple theorem, one can obtain sharp bounds for the Rayleigh quotients for symmetric matrices. This rests on work of TEMPLE [1933] and KATO [1949]; more extended work in this direction has been done by DAVIS and KAHAN [1970] (see also CHATELIN [1983, p. 46], DEMMEL [1997, Chapter 5.2], STEWART and SUN [1990, Chapter IV]).

In these results the *gap* for the ith eigenvalue plays an essential role: $gap(i, A) := \min_{j \neq i} |\lambda_j - \lambda_i|$. A small gap indicates a sensitive eigenvector. In particular, let x denote a normalized vector with Rayleigh quotient $\beta = x^T A x$, and residual $r = Ax - \beta x$. Then there is a (normalized) eigenvector q_i, corresponding to λ_i, for which

$$|\lambda_i - \beta| \leq \frac{\|r\|_2^2}{gap(i, A)}, \quad \sin(\theta) \leq \frac{\|r\|_2}{gap(i, A)}, \tag{5.4}$$

where θ denotes the acute angle between x and q_i. These results show the superior quality of a Rayleigh quotient from a given subspace. It is exploited in modern iterative methods, such as the Lanczos algorithm, but it is also essential in the QR algorithm.

Related to the perturbation analysis for Rayleigh quotients is work of KANIEL [1966] for errors in the Ritz approximations computed in the Lanczos process. For a comprehensive discussion of this, see PARLETT [1980, Chapter 12].

The theory for the effects of perturbations to eigenvectors in *unsymmetric* matrices is even more complicated and does in general not lead to easily computable error bounds. In fact, in modern analysis the attention is focussed on perturbations in invariant subspaces, see STEWART and SUN [1990, Chapter V].

We give the flavor of this for an eigenvector (invariant subspace) associated with a well isolated eigenvalue (see STEWART and SUN [1990, Chapter 5.2.3]). Let $\widetilde{A} := A+E$, x_1 is an eigenvector of A: $Ax_1 = \lambda_1 x_1$. Furthermore, $X = (x_1, Y_2)$ is a unitary matrix with first column x_1, and hence

$$X^H A X = \begin{bmatrix} \lambda_1 & G \\ 0 & A_2 \end{bmatrix}. \tag{5.5}$$

Finally, let y_1 denote the left eigenvector for λ_1 and let X_2 be defined through

$$[x_1, X_2] = [y_1, Y_2]^{-H}.$$

Then (remember that $\|x_1\|_2 = 1$, because X was assumed to be unitary)

$$\|\widetilde{x}_1 - x_1\|_2 \lesssim \|(\lambda_1 I - A)^{\#}\|_2 \|E\|_2, \tag{5.6}$$

where the so-called *Drazin inverse* is defined as

$$(\lambda_1 I - A)^{\#} := X_2(\lambda_1 I - A_2)^{-1} Y_2^H.$$

If we compare the expressions for the errors for the symmetric case (5.4) and the unsymmetric case (5.6), then we see the similarity: the factor $(\lambda_1 I - A_2)$ takes the place of the *gap*, the factors X_2 and Y_2 form a measure for the angle between x_1 and the invariant subspace X_2.

For specific perturbations one can sometimes derive less complicated error bounds. A typical situation with the eigenproblem for a large matrix $A \in \mathbb{C}^{n\times n}$ is that we have an approximate eigenpair $(\widetilde{\lambda}, \widetilde{x})$, with normalized $\widetilde{x}$, as an approximation for (λ, x), and a residual $r := (A - \widetilde{\lambda} I)\widetilde{x}$. Usually, we have not much more information than just this. If $\widetilde{\lambda} = \rho(A, \widetilde{x})$, a Rayleigh quotient, then $r^* \widetilde{x} = 0$.

Define $\widetilde{A} := A - r\widetilde{x}^*$, then

$$\widetilde{A}\widetilde{x} = \widetilde{\lambda}\widetilde{x}.$$

We can obtain an upperbound for the error in $\widetilde{\lambda}$, under the assumptions that

1. $\widetilde{A}$ can be diagonalized with X
2. r is sufficiently small

(see, for instance, Stewart and Sun [1990, Chapter IV], Ortega [1990, p. 54]).

Let X have columns $\widetilde{x}_1 = \widetilde{x}, \widetilde{x}_2, \ldots, \widetilde{x}_n$, and let the rows of X^{-1} be denoted as $s_1^{-1}\widetilde{y}_1^*$, up to $s_n^{-1}\widetilde{y}_n^*$, where

$$\|\widetilde{x}_i\|_2 = \|\widetilde{y}_j\|_2 = 1.$$

Note that the $\widetilde{y}_j$ are the left eigenvectors of $\widetilde{A}$.
We have that $s_j^{-1}\widetilde{y}_j^*\widetilde{x}_j = 1$, and hence

$$s_j = \widetilde{y}_j^*\widetilde{x}_j,$$

in particular

$$s_1 = \widetilde{y}_1^*\widetilde{x}_1. \tag{5.7}$$

Consider $\widetilde{A}$ with perturbation $\delta A := r\widetilde{x}^*$. For δA small enough, we may apply Gerschgorin's theorem, with the effect that the circle centered at $\widetilde{\lambda}$ is isolated from the other circles. This leads to

$$|\lambda - \widetilde{\lambda}| \leq \frac{1}{s_1} \sum_{j=1}^{n} |\widetilde{y}_1^* \delta A \widetilde{x}_j|. \tag{5.8}$$

Note that δA is not symmetric (we will discuss the symmetric case at the end of this section). Hence, X will in general not be orthogonal, even if A is symmetric.

Now it follows that

$$\begin{aligned} |\widetilde{y}_1^* \delta A \widetilde{x}_j| &= |\widetilde{y}_1^* r \widetilde{x}_1^* \widetilde{x}_j| \\ &= |\widetilde{y}_1^* r|\, |\widetilde{x}_1^* \widetilde{x}_j|. \end{aligned} \tag{5.9}$$

The factor $|\widetilde{y}_1^*\widetilde{x}_1|$ in (5.7) represents the cosine of the acute angle ϕ between $\widetilde{y}_1$ and $\widetilde{x}_1$:

$$\cos(\phi) = \widetilde{y}_1^*\widetilde{x}_1. \tag{5.10}$$

Hence, $\frac{\pi}{2} - \phi$ is the acute angle between $\widetilde{x}_1$ and the invariant subspace $\widetilde{y}_1^{\perp}$, spanned by $\widetilde{x}_2$, ..., $\widetilde{x}_n$, so that $|\widetilde{x}_1^*\widetilde{x}_j| \leq |\sin\phi|$. Taking this all together, we obtain an upperbound for $|\lambda - \widetilde{\lambda}|$:

$$|\lambda - \widetilde{\lambda}| \leq n\, |\widetilde{y}_1^* r|\, |\tan\phi|. \tag{5.11}$$

The factor n is pessimistic, it holds only if $\widetilde{x}_1$ makes the same small angle $\frac{\pi}{2} - \phi$ with all other $\widetilde{x}_j$.

A good practical strategy, for obtaining a realistic bound for $|\lambda - \widetilde{\lambda}|$, might be to compute the eigenvectors for a few nearby eigenvalues, determine the angles between the computed eigenvectors and the left eigenvectors of the perturbed matrix, and check whether the upper bounds (possibly with n replaced by a small integer if the angles between the computed eigenvectors are not all small) overlap. If there is no overlap then there may be some confidence in the computed eigenvalues and their bounds.

If A is symmetric, then we may choose δA as

$$\delta A := r\widetilde{x}^* + \widetilde{x}r^*, \tag{5.12}$$

which keeps $\widetilde{A}$ also symmetric. In this case X is unitary, and $\widetilde{y}_1 = \widetilde{x}_1 = \widetilde{x}$. This leads, for δA small enough (so that the Gerschgorin discs are disjunct) to

$$|\lambda - \widetilde{\lambda}| \leq \sum_{j=1}^{n} |\widetilde{x}_1^* \delta A \widetilde{x}_j| \tag{5.13}$$

and since $\widetilde{x}_1^* \widetilde{x}_j = 0$, for $j \neq 1$:

$$|\lambda - \widetilde{\lambda}| \leq \sum_{j=1}^{n} |r^* \widetilde{x}_j|. \tag{5.14}$$

Because X is unitary, the right hand side of this last inequality may be replaced by $\sqrt{n}||r||_2$.