Computing all pairs (λ, μ) such that λ is a double eigenvalue of $A + \mu B$

Elias Jarlebring,
Simen Kvaal,
Wim Michiels

Report TW 559, January 2010

Computing all pairs (λ, μ) such that λ is a double eigenvalue of $A + \mu B$

Elias Jarlebring,
Simen Kvaal,
Wim Michiels

Report TW 559, January 2010

Katholieke Universiteit Leuven
Department of Computer Science
Celestijnenlaan 200A – B-3001 Heverlee (Belgium)

Computing all pairs (λ, μ) such that λ is a double eigenvalue of $A + \mu B$

Elias Jarlebring,
Simen Kvaal,
Wim Michiels

Report TW 559, January 2010

Department of Computer Science, K.U.Leuven

Abstract

Double eigenvalues are not generic for matrices without any particular structure. A matrix depending linearly on a scalar parameter, $A + \mu B$, will however generically have double eigenvalues for some values of the parameter μ. In this paper we consider the problem of finding those values. More precisely, we construct a method to accurately find *all* scalar pairs (λ, μ) such that $A + \mu B$ has a double eigenvalue λ, where A and B are given arbitrary complex matrices.

Before presenting the numerical scheme, we prove some properties necessary for a problem to be solvable numerically in a reliable way. In particular, we show that the problem is (under mild assumptions) well conditioned.

The general idea of the globally convergent method is that if μ is close to a solution then $A + \mu B$ has two eigenvalues close to each other. We fix the relative distance between these two eigenvalues and construct a method to solve and study it by observing that the resulting problem is a *two-parameter eigenvalue problem*, which is already studied in the literature. The method, which we call the *method of fixed relative distance* (MFRD), involves solving a two-parameter eigenvalue problem which returns approximations of all solutions. It is unfortunately not possible to get full accuracy with MFRD. In order to compute solutions with full accuracy, we present an iterative method which, when given a sufficiently good starting value, returns a very accurate solution. The method returns accurate solutions for non-semisimple as well as semisimple eigenvalues.

The approach is illustrated with one academic example and one application to a simple problem in computational quantum mechanics.

COMPUTING ALL PAIRS (λ, μ) SUCH THAT λ IS A DOUBLE EIGENVALUE OF $A + \mu B$

ELIAS JARLEBRING*, SIMEN KVAAL†, AND WIM MICHIELS*

Abstract. Double eigenvalues are not generic for matrices without any particular structure. A matrix depending linearly on a scalar parameter, $A + \mu B$, will however generically have double eigenvalues for some values of the parameter μ. In this paper we consider the problem of finding those values. More precisely, we construct a method to accurately find *all* scalar pairs (λ, μ) such that $A + \mu B$ has a double eigenvalue λ, where A and B are given arbitrary complex matrices.

Before presenting the numerical scheme, we prove some properties necessary for a problem to be solvable numerically in a reliable way. In particular, we show that the problem is (under mild assumptions) well conditioned.

The general idea of the globally convergent method is that if μ is close to a solution then $A + \mu B$ has two eigenvalues close to each other. We fix the relative distance between these two eigenvalues and construct a method to solve and study it by observing that the resulting problem is a *two-parameter eigenvalue problem*, which is already studied in the literature. The method, which we call the *method of fixed relative distance* (MFRD), involves solving a two-parameter eigenvalue problem which returns approximations of all solutions. It is unfortunately not possible to get full accuracy with MFRD. In order to compute solutions with full accuracy, we present an iterative method which, when given a sufficiently good starting value, returns a very accurate solution. The method returns accurate solutions for non-semisimple as well as semisimple eigenvalues.

The approach is illustrated with one academic example and one application to a simple problem in computational quantum mechanics.

1. Introduction. For two arbitrary, given matrices $A, B \in \mathbb{C}^{n\times n}$, find the pairs $(\lambda, \mu) \in \mathbb{C}^2$ such that the matrix $A + \mu B$ has a *double* eigenvalue at λ. In this paper we will consider this problem and present a globally convergent, accurate method for all the solution pairs (λ, μ). The set of pairs (λ, μ) for which λ is a double eigenvalue of $A + \mu B$ will be denoted $\mathbb{B}(A, B)$.

Our study is in a sense a simultaneous numerical characterization of the perturbation of several double eigenvalues for a structured perturbation. Repeated eigenvalues have received a lot of attention in the field of perturbation theory. See, e.g., the standard works [Kat95, Bau85, HJ91]. Despite the fact that the understanding of repeated eigenvalues has come very far, there appears to be no globally convergent numerical method for the problem of finding all elements of $\mathbb{B}(A, B)$.

For instance, the results on generalizations of condition numbers for repeated eigenvalues [KPM09, MBO97, Sun92, Kar06, CCHI00, BLO03], is a theory for small perturbations of the repeated eigenvalue. Note that in our problem formulation, μ is generally not small, and the set $\mathbb{B}(A, B)$ is not well approximated by an asymptotic characterization for μ close to zero.

There is also recent research on the *Wilkinson distance* [AB05, Ruh70, GW76, Mai06, Mal99], which is defined as the 2-norm of the smallest perturbation E such that $A + E$ has a multiple eigenvalue. Suppose the direction of the perturbation is known, i.e., $E = \mu B$, where B is a known matrix and μ is unknown. It is easy to compute this structured Wilkinson distance with a method which computes $\mathbb{B}(A, B)$ by choosing the smallest μ in $\mathbb{B}(A, B)$. Note however, that with the methods in this paper we can compute all elements of $\mathbb{B}(A, B)$ and not only the smallest μ. Moreover, although there are estimates and numerical methods for the Wilkinson distance (see

*K.U. Leuven, Department of Computer Science, 3001 Heverlee, Belgium, {Elias.Jarlebring, Wim.Michiels}@cs.kuleuven.be

†CMA, P.O. Box 1053 Blindern, NO-0316 Oslo, Norway, simen.kvaal@cma.uio.no

e.g. the references in [AB05]), they do not seem to be easily restricted or extended to this structured case. In particular, there exist local iterative methods such as [Mai06]. The method in [Mai06] is for an unstructured problem and designed for repeated eigenvalues with an a priori known Jordan structure. There is also a method in [Mal99] which is a min-max formula for the Wilkinson distance. Note that the optimization problem is possibly difficult to solve in a reliable way. We present a method which is globally convergent and only involves linear algebra operations.

For general matrices, double eigenvalues correspond to a degenerate case and the study of double eigenvalues has traditionally been somewhat theoretically oriented. However, the results of this paper was initiated and motivated by an application. In quantum mechanical perturbation theory, the radius of convergence of a perturbation series in powers of μ for eigenvalues of $A + \mu B$ are determined by certain points in $\mathbb{B}(A, B)$ [SW73]. These points are of course not known for realistic problems, and a method to actually estimate the radius of convergence for different perturbations is of great utility. The points where there is a repeated eigenvalue is also used in the analysis of root loci in control theory. For other applications of repeated eigenvalues see the references in [Mai06].

In order to outline the contents of the paper we first briefly introduce some terminology.

Perturbations of double eigenvalues are considerably different from the theory of simple eigenvalues. For instance, the eigenvalues typically do not behave in an analytic way. Throughout this paper we will make use of the fact that λ as a function of μ is an algebraic function and can be expanded in a Puiseux series

$$\lambda(\mu + \Delta) = \lambda(\mu) + c\Delta^{\alpha} + o(\Delta^{\alpha}) \tag{1.1}$$

where α is a rational number and μ such that $\lambda(\mu)$ is a repeated eigenvalue. For double eigenvalues $\alpha \in \frac{1}{2}\mathbb{Z}$ and c can be chosen non-zero. If α is not an integer, then (1.1) characterizes a branch point of the eigenvalue map $\lambda(\mu)$, and $1/\alpha$ is called the order of this branch point.

Before presenting the numerical schemes, we prove some properties of the problem (in Section 2). It turns out that the problem is numerically well posed. We motivate this by proving preservation of existence of solutions for infinitesimal perturbations as well as continuity and conditioning properties.

It follows from the continuity of eigenvalue paths that if a matrix is a sufficiently small perturbation of a matrix having a double eigenvalue, it will have two eigenvalues close to each other. In the method presented in Section 3, we will use this property to construct a globally convergent method, based on fixing the relative distance between two eigenvalues to a small value. The resulting problem turns out to be a problem known as a *two-parameter eigenvalue problem* and can be solved with methods in the literature. We call it the *method of fixed relative distance* (MFRD).

The construction of the method in Section 3 (MFRD) is such that it will not yield an exact solution; not even in exact arithmetic. Hence, in practice, the solution will not be of full precision. Since we wish to have an accurate solution, an iterative Newton type method for elements of $\mathbb{B}(A, B)$ is presented (in Section 4). The starting value of the iterative method is taken as the result of MFRD. The combined algorithm is globally convergent and produces accurate solutions.

We illustrate the methods and theory with some examples in Section 5. We also include an Appendix providing some technical results.

2. Properties of the problem. In later sections we will construct numerical algorithms for $\mathbb{B}(A,B)$. One can only expect numerically relevant results for a problem which is posed in a numerically reasonable way. For this reason, we will in this section demonstrate some important properties of $\mathbb{B}$. Without these properties, one could not expect to solve the problem numerically. We study

- the cardinality of $\mathbb{B}(A,B)$;
- the continuous preservation of elements under perturbation; and
- the conditioning of the problem.

The conclusion of the study is that the elements of $\mathbb{B}(A,B)$ are generically finite, that the elements are always preserved under sufficiently small perturbations and that the conditioning of the problem is generically finite.

In the technical reasoning we will make use of a property that the elements if $(\lambda,\mu)\in\mathbb{B}(A,B)$ then μ is explicitly given as the root of a polynomial. To this end let

$$f(\lambda,\mu;\ A,B) := \det(\lambda I - A - B\mu). \tag{2.1}$$

A pair $(\lambda,\mu)\in\mathbb{B}(A,B)$ is now equivalently expressed as

$$\left\{\begin{array}{lcl} f(\lambda,\mu;A,B) & = & 0 \\ f_\lambda(\lambda,\mu;A,B) & = & 0. \end{array}\right. \tag{2.2}$$

Let $M_1(\mu;\ A,B)\in\mathbb{C}^{n\times n}$, respectively $M_2(\mu;\ A,B)\in\mathbb{C}^{(n-1)\times(n-1)}$ be the companion matrix corresponding to the polynomial $f(\cdot,\mu;\ A,B)$, respectively $f_\lambda(\cdot,\mu;\ A,B)$. Then we can write (2.1) as

$$\left\{\begin{array}{ll} \lambda & \in\sigma(M_1(\mu;\ A,B)) \\ -\lambda & \in\sigma(-M_2(\mu;\ A,B)), \end{array}\right. \tag{2.3}$$

which allows us to eliminate one of the variables, say λ.

PROPOSITION 2.1 (Explicit form). *Let $\mu_*\in\mathbb{C}$. The following assertions are equivalent.*

1. *$(\lambda,\mu_*)\in\mathbb{B}(A,B)$ for some $\lambda\in\mathbb{C}$.*
2. *μ_* is a solution of the polynomial eigenvalue problem*

$$\det(M_1(\mu;\ A,B)\oplus(-M_2(\mu;\ A,B))) = 0. \tag{2.4}$$

Here, we have as usual (in e.g. [HJ91]) denoted the Kronecker sum by $\oplus$. The Kronecker sum is defined as $A\oplus B = A\otimes I_1 + I_2\otimes B$, where I_1, I_2 are the identity matrices of appropriate size. We have used the property of the Kronecker sum that any eigenvalue of $A\oplus B$ is the sum of an eigenvalue of A and an eigenvalue of B.

REMARK 2.2 (Numerical stability of the explicit form). *Proposition 2.1 directly gives rise to a conceptual algorithm for the computation of $\mathbb{B}(A,B)$, which consists of first solving* (2.4) *for μ, and, next, computing the corresponding values of λ from* (2.3)*. From a numerical point of view this approach is to be avoided as it requires the explicit computation of the scalar characteristic equation, i.e., the symbolic computation of the determinant (in* (2.1)*) and the derivative.*

2.1. Cardinality. It is not reasonable to expect that a numerical algorithm can find all solutions of a problem if the problem has an infinite number of solutions. Fortunately, the generic situation is that $\mathbb{B}(A,B)$ only contains a finite number of elements.

PROPOSITION 2.3 (Cardinality). *The following classification holds.*

1. *If the nondegeneracy condition*

$$\det(M_1(\mu;\ A,B)\oplus(-M_2(\mu;\ A,B)))\not\equiv 0$$

is satisfied, then the set $\mathbb{B}(A,B)$ *consists of a finite number of isolated pairs* (λ,μ) *in* $\mathbb{R}\times\mathbb{R}$.
2. *Conversely, if*

$$\det(M_1(\mu;\ A,B)\oplus(-M_2(\mu;\ A,B)))\equiv 0, \tag{2.5}$$

then for all $\mu\in\mathbb{C}$ *there exists a* $\lambda\in\mathbb{C}$ *such that* $(\lambda,\mu)\in\mathbb{B}(A,B)$.

In this work we will focus on the first case since it is generic and the second case can be easily handled in practice by computing the eigenvalues of $A+\mu B$ for several μ. We note that the second case is sometimes referred to the case where there is a permanently double (or degenerate) eigenvalue.

2.2. Continuity. A numerical algorithm will always introduce rounding errors. If the output is not continuous with respect to the input or (more critically) solutions appear or disappear under sufficiently small perturbations, there is little hope to construct a robust numerical scheme. The problem we are considering fulfills this necessary condition, as the set $\mathbb{B}(A,B)$ is continuous with respect to changes in A and B in the following sense.

PROPOSITION 2.4 (Continuity). *Let* (λ_*,μ_*) *be an isolated pair of* $\mathbb{B}(A,B)$. *Then there exists a number* $\hat\gamma>0$ *such that for all* $\gamma\in(0,\hat\gamma)$ *there is a number* $\delta>0$ *such that* $\mathbb{B}(A+\Delta A,B+\Delta B)$ *contains at least one pair* (λ,μ) *satisfying* $|\lambda-\lambda_*|<\gamma$ *and* $|\mu-\mu_*|<\gamma$, *whenever* $\|\Delta A\|<\delta$ *and* $\|\Delta B\|<\delta$.

Proof. The proof follows from the continuous dependence of the solutions of (2.4), a one-parameter polynomial eigenvalue problem, with respect to A and B, combined with the continuous dependence of the zeros of (2.1) with respect to A,B and μ. ∎

2.3. Conditioning. Continuity with respect to changes in the input is not a sufficient for the problem to be numerically well posed. If the output is highly sensitive to perturbations in the input the problem is also considered very difficult from a numerical point of view. This comes from the fact that the first operation in a numerical algorithm will introduce rounding errors. In the following we present a characterization of the case that such ill-conditioning, i.e., high sensitivity with respect to input, occurs.

Let (λ_*,μ_*) be an isolated pair of $\mathbb{B}(A,B)$ and consider the corresponding solutions of $\mathbb{B}(A+\epsilon E_a,B+\epsilon E_b)$, where $\epsilon>0$ is a small perturbation parameter and E_a and E_b are n-by-n arbitrary complex matrices. The characterization (2.2) brings us to the equations

$$\left\{\begin{array}{lcl} f(\lambda_*+\Delta\lambda,\mu_*+\Delta\mu;A+\epsilon E_a,B+\epsilon E_b) &=& 0\\ f_\lambda(\lambda_*+\Delta\lambda,\mu_*+\Delta\mu;A+\epsilon E_a,B+\epsilon E_b) &=& 0\end{array}\right., \tag{2.6}$$

where we made the substitutions $\lambda=\lambda_*+\Delta\lambda$ and $\mu=\mu_*+\Delta\mu$. From the implicit function theorem, we conclude that if the Jacobian matrix

$$C(\lambda_*,\mu_*):=\left[\begin{array}{cc} 0 & f_\mu(\lambda_*,\mu_*;A,B)\\ f_{\lambda\lambda}(\lambda_*,\mu_*;A,B) & f_{\lambda\mu}(\lambda_*,\mu_*;A,B)\end{array}\right] \tag{2.7}$$

is invertible, then (2.6) locally defines a unique function $\epsilon\mapsto(\Delta\lambda(\epsilon),\Delta\mu(\epsilon))$ that can be expanded as

$$\begin{bmatrix} \Delta\lambda \\ \Delta\mu \end{bmatrix} = -C(\lambda_*,\mu_*)^{-1} \begin{bmatrix} \frac{\partial}{\epsilon} f(\lambda_*,\mu_*; A+\epsilon E_a, B+\epsilon E_b)\big|_{\epsilon=0} \\ \frac{\partial}{\epsilon} f_\lambda(\lambda_*,\mu_*; A+\epsilon E_a, B+\epsilon E_b)\big|_{\epsilon=0} \end{bmatrix} \epsilon + \mathcal{O}(\epsilon^2). \quad (2.8)$$

We note that the matrix $C(\lambda_*,\mu_*)$ is invertible if and only if

$$f_{\lambda\lambda}(\lambda_*,\mu_*) f_\mu(\lambda_*,\mu_*) \neq 0. \quad (2.9)$$

This condition corresponds to the generic situation where λ_* is a double, non-semisimple eigenvalue of $A+B\mu_*$ that satisfies the *completely regular splitting (CRS) property*. The local behavior of a perturbation is said to have a completely regular splitting if the order of the root in the first non-vanishing term in the Puiseux series of the eigenvalue coincide the partial multiplicity in the Jordan structure. See, e.g., [HL99] and there references therein for literature on completely regular splitting. Here we will only use the fact that CRS is the generic case. We summarize it as follows.

PROPOSITION 2.5. *Let (λ_*,μ_*) be an isolated pair of $\mathbb{B}(A,B)$.*

1. *If (2.9) holds, then the sensitivity of the pair (λ_*,μ_*) is determined by*

$$\|C(\lambda_*,\mu_*)^{-1}\|,$$

   where $C(\lambda_*,\mu_*)$ is defined by (2.7). Furthermore, the expansion (2.8) is applicable.
2. *The pair (λ_*,μ_*) is ill conditioned if $f_{\lambda\lambda}(\lambda_*,\mu_*)=0$, i.e., the eigenvalue λ_* of $A+B\mu_*$ has multiplicity larger than two.*
3. *The pair (λ_*,μ_*) is ill conditioned if $f_\mu(\lambda_*,\mu_*)=0$. This includes the case where λ^* is a double semisimple eigenvalue of $A+B\mu_*$.*

3. A method of fixed relative distance. Consider for the moment a fixed scalar $\mu \in \mathbb{C}$. Suppose $\lambda \in \mathbb{C}$ is a complex scalar fulfilling two conditions; it is an eigenvalue of $A+B\mu$, i.e.,

$$\lambda \in \sigma(A+B\mu) \quad (3.1)$$

and $(1+\varepsilon)\lambda$ is also an eigenvalue of $A+B\mu$, i.e.,

$$(1+\varepsilon)\lambda \in \sigma(A+B\mu), \quad (3.2)$$

for a fixed non-zero complex scalar $\varepsilon \in \mathbb{C}\backslash\{0\}$. The fundamental idea of the method that we will present in this section is that the solutions of (3.1) and (3.2), where λ and μ are the unknowns, approximate the elements of $\mathbb{B}(A,B)$ for sufficiently small ε. We will denote the set of all solutions of (3.1) and (3.2) by $\mathbb{D}(A,B,\varepsilon)$, i.e., $(\lambda,\mu) \in \mathbb{D}(A,B,\varepsilon)$ iff (3.1) and (3.2) hold. Note that if $(\lambda,\mu) \in \mathbb{D}(A,B,\varepsilon)$, then the matrix $A+B\mu$ has two eigenvalues with relative distance ε. The conceptual method to estimate $\mathbb{B}(A,B)$ by computing $\mathbb{D}(A,B,\varepsilon)$ for sufficiently small ε will be referred to as the *method of fixed relative distance* (MFRD).

In Proposition 2.1 we saw that if $(\cdot,\mu) \in \mathbb{B}(A,B)$, then μ were roots of a polynomial. The direct method to compute the roots of the polynomial is not attractive from a the point of view of numerical stability. This is resolved with the regularized formulation of $\mathbb{D}(A,B,\varepsilon)$. We will now see that the problem of determining $\mathbb{D}(A,B,\varepsilon)$ is a problem of a the type called *two-parameter eigenvalue problems*, which are solvable in a numerically stable way. The relations (3.1) and (3.2) can be equivalently written as a two-parameter eigenvalue problem as follows.

PROBLEM 3.1 (The associated two-parameter eigenvalue problem). *Given* $A, B \in \mathbb{C}^{n\times n}$, $\varepsilon \in \mathbb{C}$, *find* $(\lambda, \mu) \in \mathbb{C}^2$ *and a pair of non-trivial vectors* $(u, v) \in \mathbb{C}^{n\times 2}$ *such that*

$$\begin{aligned} Au &= (\lambda I - \mu B)u, \\ Av &= (\lambda(1+\varepsilon) I - \mu B)v. \end{aligned}$$

There is a general theory for the two-parameter eigenvalue problem available in the classical works [Atk72, Atk68]. There are also more recent results which are more numerically oriented [HP03], focused on the singular problem [MP09b] and generalizations [MP09a]. We also note that multiparameter eigenvalue problems have been used for (relative) placement of eigenvalues [Cot01].

There are several ways to numerically solve two-parameter eigenvalue problems, e.g., the Jacobi-Davidson type method in [HKP05]. For further methods, see the references in [HKP05]. The most common way to solve and analyze two-parameter eigenvalue problems is by means of three matrices $\Delta_0, \Delta_1, \Delta_2 \in \mathbb{C}^{n^2\times n^2}$, called the matrix determinants. In the case of Problem 3.1, the matrix determinants are,

$$\Delta_0(\varepsilon) = -I \otimes B + (1+\varepsilon) B \otimes I = ((1+\varepsilon)B) \oplus (-B), \tag{3.3}$$

$$\Delta_1 = -A \otimes B + B \otimes A, \tag{3.4}$$

$$\Delta_2(\varepsilon) = I \otimes A - (1+\varepsilon) A \otimes I = (-(1+\varepsilon)A) \oplus A. \tag{3.5}$$

One reason why the matrix determinants are important in the context of two-parameter eigenvalue problems, stems from the fact that the two-parameter eigenvalue problem is (under sufficient non-singularity conditions) equivalent to the solutions of the two coupled generalized eigenvalue problems, in this case

$$\lambda(\varepsilon)\Delta_0(\varepsilon) z = \Delta_1 z, \tag{3.6}$$

$$\mu(\varepsilon)\Delta_0(\varepsilon) z = \Delta_2(\varepsilon) z, \tag{3.7}$$

where z is a decomposable tensor $z = u \otimes v$. Although not always very efficient, we will solve the relative distance problem by solving (3.6) and (3.7). It follows from the generalized eigenvalue problems (3.6) and (3.7) that the the elements of the set $\mathbb{D}(A, B, \varepsilon)$ is (generically) a union of a finite number of functions in ε. For this reason we have denoted the solutions by $\lambda(\varepsilon)$ and $\mu(\varepsilon)$ when we wish to stress the dependence on ε.

The singularity of a two-parameter eigenvalue problem is defined as the singularity of the matrix determinant Δ_0. The non-singular two-parameter eigenvalue problem is easier to solve and analyze than the singular two-paramater eigenvalue problem. The two-parameter eigenvalue problem in Problem 3.1 is indeed non-singular for sufficiently small non-zero $\varepsilon \in \mathbb{C}\backslash\{0\}$, as can be seen in the following lemma.

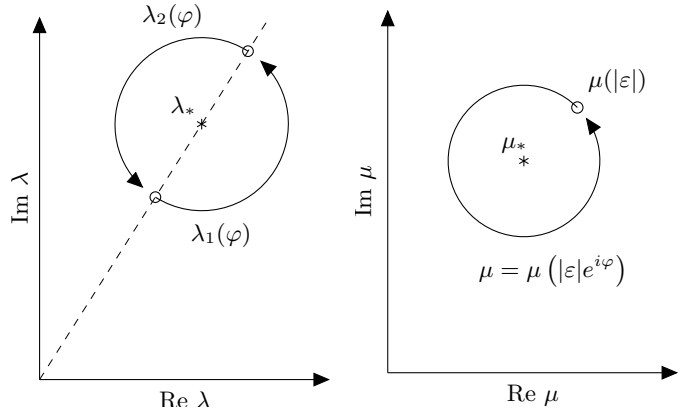
LEMMA 3.2 (Non-singularity). *The two-parameter eigenvalue problem (Problem 3.1) is singular, i.e.,* $\det(\Delta_0(\varepsilon)) = 0$, *iff there is* $\gamma \in \sigma(B)$ *such that* $\gamma(1+\varepsilon) \in \sigma(B)$.

REMARK 3.3 (The non-zero regularization parameter ε). *If* ε *is set to zero, it is easy to see from Lemma 3.2 that the two-parameter eigenvalue problem is singular. Although there are results on singular two-parameter eigenvalue problems [MP09b] such an approach seems impossible for this problem. Note that if* $\varepsilon = 0$, *then the two eigenvalue problems* (3.1) *and* (3.2) *are identical and the solution corresponds to the eigenvalue curves* $\lambda(\mu) \in \sigma(A + \mu B)$, *which do not correspond to double eigenvalues.*

The two-parameter eigenvalue problem for $\varepsilon = 0$ is singular and the corresponding problem for $\mathbb{D}(A, B, \varepsilon)$ is not singular. Hence, the problem associated with MFRD is in a sense a regularized problem and we will call the parameter ε a regularization parameter.

In the remaining parts of this section we wish to show properties of $\mathbb{D}(A, B, \varepsilon)$ and in particular in what sense $\mathbb{D}(A, B, \varepsilon)$ approximates $\mathbb{B}(A, B)$.

3.1. Consistency. The first property which we will illustrate in a remark and a theorem is that the method is consistent, in the sense that (generically) all finite limits of $\mathbb{D}(A, B, \varepsilon)$ as $\varepsilon \to 0$ coincide with $\mathbb{B}(A, B)$. Moreover, this limit is independent of the angle of ε.



(a) $\lambda(\varepsilon)$ runs $\frac{1}{2}$ lap (b) $\mu(\varepsilon)$ runs one lap

FIGURE 3.1. *The approximations rotate around the solutions λ_* and μ_* as angle of ε is changed. The plots show the curves for $\varphi \in [0, \pi]$. We have denoted $\lambda_1(\varphi) := \lambda\left(|\varepsilon|e^{i\varphi}\right)$ and $\lambda_2(\varphi) := \lambda_1(\varphi)(1 + |\varepsilon|e^{i\varphi})$. The dashed line corresponds to $\varphi = 0$. In subfigure (a), the same dashed line (asymptotically) corresponds to the line to the origin.*

REMARK 3.4 (The angle of ε). *Since ε is the relative distance, fixing the complex angle of ε fixes the angle of the separation of the eigenvalues $\lambda(\varepsilon)$ and $(1 + \varepsilon)\lambda(\varepsilon)$. We will now see that the angle of ε will asymptotically only influence the angle of the approximation error. The asymptotic effect of changing φ where $\varepsilon = |\varepsilon|e^{i\varphi}$ can be roughly motivated as follows. Consider the first terms in a completely regular square root splitting, $\lambda(\mu(\varepsilon)) = \lambda_* \pm \sqrt{\mu_* - \mu(\varepsilon)}$. The condition that the relative distance is ε implies that $\mu_* - \mu(\varepsilon) = c\varepsilon^2 = (2c)^{-1/2}|\varepsilon|^2 e^{i2\varphi}$. This in turn implies that $\lambda(\varepsilon) - \lambda_* = \pm\frac{1}{2}|\varepsilon|e^{\pm i\varphi}$. That is, φ only influences the angle of the error. The phenomenon is illustrated in Figure 3.1. The same reasoning reasoning holds for an arbitrary splitting of the eigenvalue.*

In the following theorem we formalize the argment above and see that the finite limits of $\mathbb{D}(A, B, \varepsilon)$ as $\varepsilon \to 0$ are generically equal to the set $\mathbb{B}(A, B)$.

THEOREM 3.5 (Consistency). *Consider the pair $A, B \in \mathbb{C}^{n\times n}$. The sets $\mathbb{D}(A, B, \varepsilon)$ and $\mathbb{B}(A, B)$ are related by the following statements:*

i) If B is non-singular, the set $\mathbb{D}(A, B, \varepsilon)$ is the union of n^2 pair of functions $(\lambda(\varepsilon), \mu(\varepsilon))$ for sufficiently small $\varepsilon \neq 0$.

ii) If $(\lambda(\varepsilon), \mu(\varepsilon)) \in \mathbb{D}(A, B, \varepsilon)$ *and* $(\lambda(\varepsilon), \mu(\varepsilon)) \to (\lambda_*, \mu_*) \in \mathbb{C}^2$ *as* $|\varepsilon| \to 0$ *and* $\lambda_* \neq 0$ *then* $(\lambda_*, \mu_*) \in \mathbb{B}(A, B)$.

iii) If $(\lambda_*, \mu_*) \in \mathbb{B}(A, B)$ *and* $\lambda_* \neq 0$ *then there exists a path* $(\lambda(\varepsilon), \mu(\varepsilon)) \in \mathbb{D}(A, B, \varepsilon)$ *such that* $(\lambda(\varepsilon), \mu(\varepsilon)) \to (\lambda_*, \mu_*)$ *as* $|\varepsilon| \to 0$.

Proof. The paths $\lambda(\varepsilon)$ and $\mu(\varepsilon)$ are solutions of the generalized eigenvalue problems corresponding to the pencils $\Delta_1 - \lambda(\varepsilon)\Delta_0(\varepsilon)$ and $\Delta_2(\varepsilon) - \mu(\varepsilon)\Delta_0(\varepsilon)$. Eigenvalue problems depending on continuously a parameter is the union of paths, where the number of paths equals the dimension of the matrix, here n^2 (see e.g. [HP05, Corollary 4.2.4]). The generalized eigenvalue problems can be rewritten as standard eigenvalue problems if $\Delta_0(\varepsilon)$ is non-singular. Hence, to show i) it is suffient to show that $\Delta_0(\varepsilon)$ is non-singular for sufficiently small ε.

It follows from the sum-property of Kronecker sums that the eigenvalues of $\Delta_0(\varepsilon)$ are $(1+\varepsilon)b_i - b_j$ for $i, j = 0, \dots n$ where b_i are the eigenvalues of B. Since $b_i \neq 0$, all eigenvalues of $\Delta_0(\varepsilon)$ are non-zero for sufficiently small $\varepsilon \neq 0$. This proves i).

Let $f(\lambda, \mu) = \det(-\lambda I + A + B\mu)$. In order to show ii) we will show that $f_\lambda(\lambda_*, \mu_*) = 0$. First note that

$$f(\lambda(\varepsilon), \mu(\varepsilon)) = 0,$$
$$f(\lambda(\varepsilon)(1+\varepsilon), \mu(\varepsilon)) = 0.$$

Hence, by Taylor expansion,

$$0 = f(\lambda(\varepsilon)(1+\varepsilon), \mu(\varepsilon)) = f(\lambda(\varepsilon), \mu(\varepsilon)) + \lambda(\varepsilon)\varepsilon f_\lambda(\lambda(\varepsilon), \mu(\varepsilon)) + \mathcal{O}(\lambda(\varepsilon)\varepsilon)^2.$$

Since $\lambda(\varepsilon) \to \lambda_* \in \mathbb{C}\backslash\{0\}$ when $\varepsilon \to 0$, it holds that $f_\lambda(\lambda_*, \mu_*)$. ∎

REMARK 3.6 ($\lambda = 0$). *Note that the case* $\lambda = 0$ *has to be explicitly excluded in Theorem 3.5. This stems from the fact that the relative distance is defined from* (3.1) *and* (3.2)*, which are trivially fulfilled if* $\lambda = 0$ *(and* μ *an eigenvalue of the generalized eigenvalue problem* $A + \mu B$*) for any* ε*. The special case that* $\lambda = 0$ *is a double eigenvalue can be easily handled by hand and the solutions corresponding to* λ *can be safely and easily excluded from the solution in the implementation.*

3.2. Unbounded limits. We now know (from Theorem 3.5) that the finite limits of $\mathbb{D}(A, B, \varepsilon)$ form in a relevant way the set $\mathbb{B}(A, B)$. At this point, it is important to note that this does not necessarily imply that the elements of $\mathbb{D}(A, B, \varepsilon)$ always approximate a corresponding element of $\mathbb{B}(A, B)$ as $\varepsilon \to 0$. It turns out that some elements of $\mathbb{D}(A, B, \varepsilon)$ can be unbounded as $\varepsilon \to 0$. The cases where this occurs can fortunately be completely identified before the method start as they are completely characterized by the multiplicity of the eigenvalues of B.

In order to show this we need the following result which relates $\mathbb{D}(A, B, \varepsilon)$ with the converse problem $\mathbb{D}(B, A, \varepsilon)$.

LEMMA 3.7 (A converse identity). *Let* $\varepsilon, \mu \in \mathbb{C} \setminus \{0\}$ *be given. Then,*

$$(\lambda, \mu) \in \mathbb{D}(A, B, \varepsilon) \;\Leftrightarrow\; (\lambda/\mu, 1/\mu) \in \mathbb{D}(B, A, \varepsilon).$$

Proof. Define $f(\lambda, \mu) = \det(A + \mu B - \lambda I)$ and $g(\gamma, \nu) = \det(B + \nu A - \gamma I)$. Now, $(\lambda, \mu) \in \mathbb{D}(A, B, \varepsilon)$ if and only if $f(\lambda, \mu) = f((1+\varepsilon)\lambda, \mu) = 0$ if and only if $g(\lambda/\mu, 1/\mu) = g((1+\varepsilon)\lambda/\mu, 1/\mu) = 0$, since $f(\mu, \lambda) = \mu^n g(\lambda/\mu, 1/\mu)$. Since the relative distance between λ/μ and $(1+\varepsilon)\lambda/\mu$ is ε, $(\lambda/\mu, 1/\mu) \in \mathbb{D}(B, A, \varepsilon)$ as claimed. ∎

In the following theorem we see a characterization of those elements of $\mathbb{D}(A, B, \varepsilon)$ which are unbounded as $\varepsilon \to 0$. An element is unbounded in magnitude as $\varepsilon \to 0$ if the matrix B has a repeated eigenvalue.

THEOREM 3.8 (Unbounded elements). *Assume that $A + \mu B$ has no permanently degenerate eigenvalues and that B is invertible. Let $\gamma_* \in \mathbb{C} \setminus \{0\}$. Then these statements are equivalent:*

1. *γ_* is a multiple eigenvalue of B*
2. *There exists a punctured disc $\mathcal{D} = \{\varepsilon : |\varepsilon| < \hat{\varepsilon}\} \setminus \{0\} \subset \mathbb{C}$ and functions*

$$\varepsilon \in \mathcal{D} \mapsto (\lambda(\varepsilon), \mu(\varepsilon)) \in \mathbb{C} \times \mathbb{C}$$

satisfying $(\lambda(\varepsilon), \mu(\varepsilon)) \in \mathbb{D}(A, B, \varepsilon)$ for all $\varepsilon \in \mathcal{D}$ with

$$\lim_{\varepsilon \to 0} |\lambda(\varepsilon)| = \infty$$
$$\lim_{\varepsilon \to 0} \frac{\lambda(\varepsilon)}{\mu(\varepsilon)} = \gamma_*.$$

Proof. Let $f(\lambda, \mu)$ and $g(\gamma, \nu)$ be as in Lemma 3.7. We first prove that 1 implies 2. Since B is invertible, $\gamma_* \neq 0$. By point *(iii)* of Theorem 3.5, there exist a small puncured disk $\mathcal{D}$ and functions $\gamma, \nu : \mathcal{D} \to \mathbb{C} \times \mathbb{C}$ with $(\gamma(\varepsilon), \nu(\varepsilon)) \in \mathbb{D}(B, A, \varepsilon)$ and $\lim_{\varepsilon \to 0}(\gamma(\varepsilon), \nu(\varepsilon)) = (\gamma_*, 0)$.

By Lemma 3.7,

$$(\lambda(\varepsilon), \mu(\varepsilon)) := \left(\frac{\gamma(\varepsilon)}{\nu(\varepsilon)}, \frac{1}{\nu(\varepsilon)}\right) \in \mathbb{D}(A, B, \varepsilon) \quad \forall \varepsilon \in \mathcal{D}.$$

Moreover,

$$\lim_{\varepsilon \to 0} |\lambda(\varepsilon)| = \lim_{\varepsilon \to 0} \left|\frac{\gamma(\varepsilon)}{\nu(\varepsilon)}\right| = \infty$$

since $\nu(\varepsilon) \to 0$ and $\gamma(\varepsilon) \to \gamma_* \neq 0$. We also have

$$\lim_{\varepsilon \to 0} \frac{\lambda(\varepsilon)}{\mu(\varepsilon)} = \lim_{\varepsilon \to 0} \gamma(\varepsilon) = \gamma_*.$$

The implication is thus proven.

To prove that 2 implies 1, we suppose that the functions $\lambda, \mu : \mathcal{D} \to \mathbb{C} \times \mathbb{C}$ exists, with the desired limits. Define $\gamma(\varepsilon) := \lambda(\varepsilon)/\mu(\varepsilon)$ and $\nu(\varepsilon) := 1/\mu(\varepsilon)$. By Lemma 3.7, $(\gamma(\varepsilon), \nu(\varepsilon)) \in \mathbb{D}(B, A, \varepsilon)$. Since $\lambda(\varepsilon) \in \sigma(A + B\mu(\varepsilon))$ and $\lambda(\varepsilon) \to \infty$ we must also have $\mu(\varepsilon) \to \infty$ as $\varepsilon \to 0$.

Consequently,

$$\lim_{\varepsilon \to 0} \nu(\varepsilon) = \lim_{\varepsilon \to 0} \frac{1}{\mu(\varepsilon)} = 0,$$

and $(\gamma(\varepsilon), \nu(\varepsilon)) \to (\gamma_*, 0)$ as $\varepsilon \to 0$. By point *(ii)* of Theorem 3.5, γ_* is then a double eigenvalue of B. This concludes the proof. ∎

3.3. Regularization error. Since ε must be chosen non-zero and not too small in practice (see Remark 3.3), the construction with relative distance will always generate some error, which we will call the regularization error. In this section we will

see that this error is reasonably behaved. We show this by proving some results about the asymptotic error as a function of the regularization parameter ε.

If the repeated eigenvalue has a completely regular square root splitting, which is the generic case, the accuracies are $\lambda(\varepsilon)-\lambda_* = \mathcal{O}(\varepsilon)$, i.e., linear, and $\mu(\varepsilon)-\mu_* = \mathcal{O}(\varepsilon^2)$, i.e., quadratic.

We will later use this result to propose a rough argument for how to choose ε in practice.

THEOREM 3.9 (Regularization error). *Let (λ_*, μ_*) be an isolated pair in $\mathbb{B}(A, B)$, then there exists a pair of functions $(\lambda(\varepsilon), \mu(\varepsilon)) \in \mathbb{D}(A, B, \varepsilon)$ such that*

$$\lim_{\varepsilon\to 0} \frac{|\lambda(\varepsilon) - \lambda_*|}{|\varepsilon|^\alpha} = C_1|\lambda_*|^\alpha, \tag{3.8}$$

and

$$\lim_{\varepsilon\to 0} \frac{|\mu(\varepsilon) - \mu_*|}{|\varepsilon|^\beta} = C_2|\lambda_*|, \tag{3.9}$$

where $C_1, C_2 \in \mathbb{R}$ and $\alpha, \beta \in \mathbb{R}_+$ are independent of ε. Moreover, if $\lambda(\cdot)$ has a completely regular square root splitting with $\lambda(\varepsilon) = \lambda_* \pm C(\mu_* - \mu(\varepsilon))^{1/2} + o(\mu_* - \mu(\varepsilon))$ with $C \neq 0$, then the constants are $\alpha = 1$, $\beta = 2$, $C_1 = \frac{1}{2}$ and $C_2 = \frac{1}{|4C^2|}$.

Proof. The proof for (3.8) and (3.9) is done for the general setting where the eigenvalue $\lambda(\varepsilon)$ has an Puiseux expansion,

$$\lambda(\varepsilon) = \lambda_* + C(\mu(\varepsilon) - \mu_*)^q + o(\mu(\varepsilon) - \mu_*)^q, \tag{3.10}$$

and $\lambda_2(\varepsilon)$ is either a different branch from the same Puiseux series or a different Puiseux expansion. First consider the Puiseux expansion of the difference,

$$\lambda(\varepsilon) - \lambda_2(\varepsilon) = D(\mu(\varepsilon) - \mu_*)^p + o(\mu(\varepsilon) - \mu_*)^p,$$

where C,$D \in \mathbb{C}$. Note that C and D can be chosen non-zero since the eigenvalue is not permanently double. The two equations can now be solved by eliminating $(\mu(\varepsilon) - \mu_*)^{qp}$,

$$\begin{aligned}&((\lambda(\varepsilon) - \lambda_*)^p/C^p + o(\mu(\varepsilon) - \mu_*)^q)^p = (\mu(\varepsilon) - \mu_*)^{qp} = \\ &(\lambda(\varepsilon) - \lambda_2(\varepsilon))^q/D^q + o(\mu(\varepsilon) - \mu_*)^p)^q.\end{aligned}$$

The relative distance is ε by construction, i.e., $\lambda_2(\varepsilon) = \lambda(\varepsilon)(1 + \varepsilon)$. Inserting $\lambda_2(\varepsilon) = \lambda(\varepsilon)(1 + \varepsilon)$ into the second equation, yields that

$$|\lambda(\varepsilon) - \lambda_*|^p/|C|^p = |\varepsilon|^q|\lambda(\varepsilon)|^q/|D|^q + o(\mu(\varepsilon) - \mu_*)^p.$$

Hence,

$$\frac{|\lambda(\varepsilon) - \lambda_*|}{|\varepsilon|^{q/p}} = \frac{|C|}{|D|^{q/p}}|\lambda(\varepsilon)|^{q/p} + o(\mu(\varepsilon) - \mu_*).$$

This proves (3.8). From (3.10) we find that

$$|\mu(\varepsilon) - \mu_*| = \frac{|\lambda(\varepsilon)||\varepsilon|^{1/p}}{|D|^{1/p}} + o(|\varepsilon|^{1/p}).$$

We have proven (3.9). For a completely regular square root splitting we can choose $p = q = 1/2$ and $D = 2C$. Hence, $\alpha = 1$, $\beta = 1/p = 2$, $C_1 = |C/2C| = 1/2$, $C_2 = 1/(4C^2)$. ∎

REMARK 3.10 (Extrapolation). *The algebraic convergence of the solution as $\varepsilon \to 0$ can be used to accelerate convergence by evaluating for several ε. This extrapolation technique will not be used in this work since evaluating $(\lambda(\varepsilon), \mu(\varepsilon))$ is the most computationally dominating part of the method. In order to gain full accuracy we will (in Section 4) instead use a local iterative method to gain high accuracy.*

3.4. The regularization parameter trade-off. In practice we must fix the regularization parameter ε before the method starts. If ε is chosen too large, the regularization error will (according to Theorem 3.9) be large. If ε is chosen too small it is natural to expect that rounding errors will destroy the accuracy of the solution since the problem is very close to singular. We now wish to provide a rough estimate of ε to find a reasonably good resolution to the trade-off.

In this subsection we will restrict the error analysis for the matrix determinant approach to solve the two-parameter eigenvalue problem, i.e., we will solve the generalized eigenvalue problem corresponding to the pencil

$$D_\varepsilon(\lambda) := \Delta_1 - \lambda \Delta_0(\varepsilon).$$

We now wish to outline in what way the magnitude of ε affects the condition of the eigenvalue problem of D_ε for $\varepsilon \neq 0$. This will lead us to a reasoned choice of ε to be used in the implementation.

We start with a technical lemma that shows that the structure of the problem is such that the first terms in the expansion of $\det(D_\varepsilon(\lambda))$ (in powers of ε) vanish.

LEMMA 3.11. *The following expansion holds:*

$$\det D_\varepsilon(\lambda) = \sum_{k=2}^{n^2} \varepsilon^k f_k(\lambda),$$

where the functions f_k, $k = 2, \ldots, n^2$, are polynomials of degree smaller than or equal to n^2.

Proof. By explicitly computing the determinant it follows that

$$\det D_\varepsilon(\lambda) = \sum_{k=0}^{n^2} \varepsilon^k f_k(\lambda),$$

where f_k, $k = 0, \ldots, n^2$, are polynomials. Due to the fact that the pencil $D_0(\lambda)$ is singular, we have $f_0 \equiv 0$. It remains to prove that $f_1 \equiv 0$.

For an arbitrary fixed value of λ, we get

$$f_1(\lambda) \;\; = \left. \tfrac{\partial \det D_\varepsilon(\lambda)}{\partial \varepsilon} \right|_{\varepsilon=0} = \mathrm{Tr} \, \left\{ \mathrm{adj}(D_0(\lambda)) \left. \tfrac{\partial D_\varepsilon(\lambda)}{\partial \varepsilon} \right|_{\varepsilon=0} \right\}.$$

We can express

$$\begin{aligned} D_0(\lambda) &= (-A \otimes B + B \otimes A) - \lambda(-I \otimes B + B \otimes I) \\ &= B \otimes (A - \lambda I) - (A - \lambda I) \otimes B. \end{aligned}$$

If $(A - \lambda I)$ is regular, then we get from Lemma A.1 in the appendix that $\mathrm{adj}\, D_0(\lambda) = 0$. We conclude that $f_1(\lambda) = 0$ for all $\lambda \notin \sigma(A)$. Because f_1 is a polynomial this implies that $f_1 \equiv 0$ and the proof is completed. ∎

In what follows we investigate the condition of the eigenvalue problem of the pencil D_ε, for a fixed value of $\varepsilon \neq 0$. In order to assess the effect of perturbations of the matrices Δ_0 and Δ_1 on the eigenvalues of D_ε we consider the pseudospectra $\Lambda_\gamma(D_\varepsilon)$, $\gamma > 0$, defined as:

$$\Lambda_\gamma(D_\varepsilon) := \{\lambda \in \mathbb{C} : \ \det\{(\Delta_1 + \delta\Delta_1) - \lambda(\Delta_0(\varepsilon) + \delta\Delta_0)\} = 0 \text{ for some } \delta\Delta_0, \delta\Delta_1 \in \mathbb{C}^{n^2\times n^2}, \text{ satisfying } \frac{\|\delta\Delta_0(\varepsilon)\|_2}{\|\Delta_0\|_2} < \gamma \text{ and } \frac{\|\delta\Delta_1\|_2}{\|\Delta_1\|_2} < \gamma\Big\}. \quad (3.11)$$

Thus, the pseudospectrum $\Lambda_\gamma(D_\varepsilon)$ is a subset of the complex plane consisting of all possible positions of the eigenvalues of D_ε when the system matrices are subjected to perturbations with relative size smaller than γ. From Theorem 1 of [MGWN06] the following computational formula can be derived:

$$\Lambda_\gamma(D_\varepsilon) = \left\{\lambda \in \mathbb{C} : \ \|(D_\varepsilon(\lambda))^{-1}\|_2(\|\Delta_0(\varepsilon)\|_2 + |\lambda|\|\Delta_1\|_2) > \frac{1}{\delta}\right\}. \quad (3.12)$$

Now, let λ_ε be an isolated eigenvalue of D_ε, that is,

$$\det D_\varepsilon(\lambda_\varepsilon) = 0, \quad \frac{d}{d\lambda}\det D_\varepsilon(\lambda)\Big|_{\lambda=\lambda_\varepsilon} \neq 0.$$

If $|\lambda - \lambda_\varepsilon|$ is small we can approximate:

$$\|(D_\varepsilon(\lambda))^{-1}\|_2(\|\Delta_0(\varepsilon)\|_2 + |\lambda|\|\Delta_1\|_2) \approx \frac{\|\text{adj } D_\varepsilon(\lambda_\varepsilon)\|_2}{\left|\frac{d}{d\lambda}\det D_\varepsilon(\lambda)\big|_{\lambda=\lambda_\varepsilon}\right| |(\lambda - \lambda_\varepsilon)|}(\|\Delta_0(\varepsilon)\|_2 + |\lambda_\varepsilon|\|\Delta_1\|_2). \quad (3.13)$$

From (3.12) and 3.13 we conclude that for sufficiently small values of γ, the pseudospectrum Λ_γ contains a disc centered around λ_ε, with radius equal to $r(\lambda_\varepsilon)\delta$, where

$$r(\lambda_\varepsilon) := \frac{\|\text{adj } D_\varepsilon(\lambda_\varepsilon)\|_2}{\left|\frac{d}{d\lambda}\det D_\varepsilon(\lambda)\big|_{\lambda=\lambda_\varepsilon}\right|}(\|\Delta_0(\varepsilon)\|_2 + |\lambda_\varepsilon|\|\Delta_1\|_2). \quad (3.14)$$

In words, the number $r(\lambda_\varepsilon)$ is the growth rate of the pseudospectrum $\Lambda_\gamma(D_\varepsilon)$ around the eigenvalue λ_ε when δ is increased from zero.

REMARK 3.12. *The number $r(\lambda_\varepsilon)$ corresponds to the (structured)* condition number *for the eigenvalue λ_ε as defined in [ARK08, Equation (4)]. When using the property* adj $D_\varepsilon(\lambda_\varepsilon) = c\, uv^*$, *where u and v are normalized right and left null vectors of $D_\varepsilon(\lambda_\varepsilon)$ and $c \in \mathbb{C}$, the expression* (3.14) *can be simplified to*

$$r(\lambda_\varepsilon) = \frac{1}{v^*u}(\|\Delta_0(\varepsilon)\|_2 + |\lambda_\varepsilon|\|\Delta_1\|_2),$$

which is consistent with the expression formulated in [ARK08, Lemma 2.1].

Taking into account Lemma 3.11 we can simplify (3.14) to

$$r(\lambda_\varepsilon) = \frac{1}{\varepsilon^2}\frac{\|\text{adj } D_\varepsilon(\lambda_\varepsilon)\|_2}{\left|\sum_{k=2}^{n^2} f_k'(\lambda_\varepsilon)\varepsilon^{k-2})\right|}(\|\Delta_0(\varepsilon)\|_2 + |\lambda_\varepsilon|\|\Delta_1(\lambda_\varepsilon)\|_2).$$

Hence, if $\lim_{\varepsilon\to 0}\lambda_\varepsilon$ is finite, say λ_*, then the pseudospectral growth rate increases inversely proportional to ε^2 as $\varepsilon \to 0$. We can now come to the main point of this section. On the one hand, if we apply a stable algorithm to compute the eigenvalues of D_ε (for instance, the celebrated QZ algorithm) it is expected that the computational error on the result is comparable to the error induced by rounding errors on the data. Therefore, for a fixed value of ε the worst-case computational error on the eigenvalue λ_ε is expected to be proportional to

$$E_c(\varepsilon) := r(\lambda_\varepsilon)\ \varepsilon_{\text{mach}} \sim \frac{\varepsilon_{\text{mach}}}{\varepsilon^2}, \tag{3.15}$$

where $\varepsilon_{\text{mach}}$ is the machine precision. On the other hand, recall from Theorem 3.9 that the approximation error, that is $E_a := (\lambda(\varepsilon) - \lambda_*)/\lambda_*$, satisfies

$$E_a(\varepsilon) \sim \varepsilon, \tag{3.16}$$

which needs to be small, in order to obtain good approximations of all solutions of $\mathbb{B}(A, B)$. The optimal choice of ε involves a trade-off between (3.15) and (3.16). It leads us to the choice

$$\varepsilon \sim \varepsilon_{\text{mach}}^{1/3}, \tag{3.17}$$

for which both

$$E_c \sim \varepsilon_{\text{mach}}^{1/3}, \quad E_a \sim \varepsilon_{\text{mach}}^{1/3}.$$

As demonstrated in Section 5, this must be treated as a rule of thumb. Note also that a similar reasoning can be done for the pencil $\Delta_2(\varepsilon) - \mu\Delta_0(\varepsilon)$.

4. Local methods. The method in Section 3 (MFRD) has the attractive property that it approximates all solutions of $\mathbb{B}(A, B)$. A drawback of the method is that a regularization error is introduced dependent on the parameter ε which can not be chosen zero or too small. In this section, we show how, in a post-processing step, approximations can be improved using local iterative methods. These methods are based on solving systems of nonlinear equations that fully characterize the elements of $\mathbb{B}(A, B)$. Since they are iterative methods, they rely on starting values, which can be generated from the solutions of $\mathbb{D}(A, B, \varepsilon)$.

A double eigenvalue can either be semisimple or non-semisimple. If it is semisimple then

$$\begin{aligned}(A + B\mu - \lambda I)v_1 &= 0\\ (A + B\mu - \lambda I)v_2 &= 0\end{aligned}$$

for non-parallel v_1 and v_2 whereas for non-semisimple eigenvalues, we have a generalized eigenvector u associated with the eigenvector v_1, such that

$$\begin{aligned}(A + B\mu - \lambda I)v &= 0\\ (A + B\mu - \lambda I)u &= v.\end{aligned}$$

We do not know a priori which case occurs. In order to construct a local iterative methods which works for both semisimple and non-semisimple eigenvalues we will consider the null space of $(A + B\mu - \lambda I)^2$. We will use the following property (which

is also used in [Mal99, Lemma 1]). It is easy to see that $(A + B\mu - \lambda I)^2$ has a two-dimensional null space independent of the two cases semisimple and non-semisimple.

In this way we arrive at the nonlinear equations

$$\begin{cases} (A + B\mu - \lambda I)^2 v_1 = 0 \\ (A + B\mu - \lambda I)^2 v_2 = 0 \\ a_1^* v_1 = 1 \\ a_2^* v_1 = 1 \\ a_1^* v_2 = 1 \\ v_1^* v_2 = 0 \end{cases} \tag{4.1}$$

in the unknowns (λ, μ, v_1, v_2). Note that the four last equations, with $a_1 \neq a_2$, are normalizing conditions that allow to specify uniquely a basis of a two-dimensional null space. With starting values obtained form the elements of $\mathbb{D}(A, B, \varepsilon)$ we can solve the overdetermined system (4.1) using the Gauss-Newton method.

REMARK 4.1 (Relation with Newton's method). *When considering only the first, the third and the fourth equation the system becomes fully determined and can be solved using the standard Newton's method. The underlying idea is as follows. By means of the third and fourth equation we require that the null vector v_1 satisfies* two *normalization constraints simultaneously. If the null space is one-dimensional (which implies $(\lambda, \mu) \notin \mathbb{B}(A, B)$) it is generically not possible to satisfy both normalization constraints, unlike in the case where the null space is at least two-dimensional (which implies $(\lambda, \mu) \in \mathbb{B}(A, B)$). The motivation for keeping the other equations is that they create an additional penalty (residual) when the iterate deviates from the desired solution. This leads to a better conditioned problem and a more accurate numerical solution, as confirmed by our experiments.*

REMARK 4.2. *In the case where λ_* is a* non-semisimple *eigenvalue of $A+B\mu_*$, an alternative method consists of applying the algorithm of [Mai06]. The latter is based on an application of Newton's method to solve a system of equations that characterizes the presence of a Jordan block of at least two-by-two in the canonical representation. In our experiments it yields a performance comparable to solving* (4.1) *with Gauss-Newton.*

In Section B of the appendix the following is shown: if the solution of (4.1) corresponds to a double non-semisimple eigenvalue satisfying the completely regular splitting property, then the problem of solving (4.1) is well conditioned, in the sense that the Jacobian matrix in the solution generically has full column rank. The latter further implies that the Gauss-Newton algorithm exhibits quadratic convergence.

In the case where the double eigenvalue is *semi-simple*, the problem of solving (4.1) is ill conditioned due a rank deficient Jacobian in the solution. This affects the accuracy of the obtained numerical results. As outlined also in the Appendix, this problem can be overcome by instead solving the equations

$$\begin{cases} (A + B\mu - \lambda I) v_1 = 0 \\ (A + B\mu - \lambda I) v_2 = 0 \\ a_1^* v_1 = 1 \\ a_2^* v_1 = 1 \\ a_1^* v_2 = 1 \\ v_1^* v_2 = 0 \end{cases}, \tag{4.2}$$

which directly characterize the presence of a two-dimensional eigenspace of the matrix $A + B\mu$.

In our implementation we start by solving (4.1). If the resulting matrix $A+B\mu-\lambda I$ has *two* singular values below a threshold (10^{-6} in our code), we consider this as is an indication that the eigenvalue could be semi-simple and the algorithm attempts solving (4.2) in a subsequent step, with starting values obtained from the solution of (4.1). In case of convergence we conclude that the eigenvalue is semi-simple.

5. Examples.

5.1. A 3×3 test problem. Suppose

$$A=\begin{pmatrix}-1 & 2 & 1\\ 0 & 2 & -\mathrm{i}\\ \mathrm{i} & 1 & -\mathrm{i}\end{pmatrix},\quad B=(C-A)/\mu_0,\quad C=\begin{pmatrix}1 & 0 & 0\\ 0 & 2 & 0\\ 0 & 0 & 2\end{pmatrix}, \tag{5.1}$$

where $\mu_0 = 1+\mathrm{i}$. This problem is constructed such that for $\mu=\mu_0$, $\lambda=2$ is a semi-simple eigenvalue of $A+\mu B$. It can be solved explicitly with software for symbolic manipulations by simultaneously solving $f_\lambda(\lambda,\mu)=0$ and $f(\lambda,\mu)=0$ as in Remark 2.2. The solution is $\mathbb{B}(A,B)\approx\{(1+i,2),(0.60+0.40i,0.50-0.39i),(0.98+1.4i,1.6+0.32),(1.1+1.3i,2.1+0.21i),(1.5+1.2i,1.9-0.2i)\}$. We will use this solution (with sufficiently high precision) for reference.

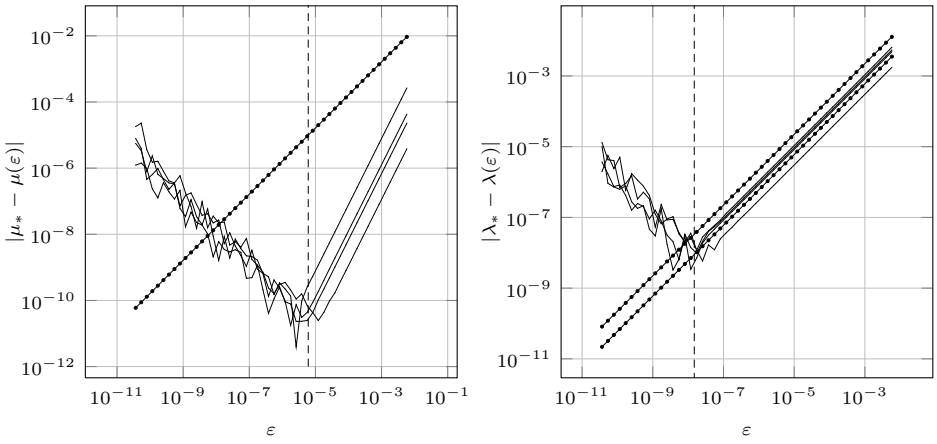
This example will now be used to illustrate how the accuracy of the solution computed with MFRD depends on the regularization parameter ε. The error for the solutions are given in Figure 5.1. We observe V-shaped error curves (as predicted in Section 3.4) for the non-semisimple eigenvalues, corresponding essentially to the maximum of the rounding error (dominating for small ε) and regularization error (dominating for larger ε).

Note also that there is no optimal choice ε in the sense that the errors of individual approximation will never be simultaneously minimized. This holds in particular for the semi-simple eigenvalue for which the regularization error is of order $|\mu-\mu(\varepsilon)|=O(\varepsilon)$ (unlike the other error curves $|\mu-\mu(\varepsilon)|=O(\varepsilon^2)$). Hence, there is no choice of ε such that MFRD returns full precision solutions. This problem is not present when we combine MFRD with the individual treatment of the approximations with the iterative method in Section 4. The iterative method computes solutions to an accuracy of a small multiple times $\varepsilon_{\mathrm{mach}}$ for all elements of $\mathbb{B}(A,B)$.

5.2. Application to quantum mechanical perturbation theory. An important problem in computational quantum mechanics is to characterize the smallest eigenvalue (assuming this exists) of a self-adjoint partial differential operator $\mathcal{H}=\mathcal{A}+\mu_{\mathrm{phys}}\mathcal{B}$ (with $\mu_{\mathrm{phys}}\in\mathbb{R}$) over some infinite dimensional Hilbert space. Standard discretization techniques produce a (Hermitean) matrix approximation $H=A+\mu_{\mathrm{phys}}B$ to this operator. The matrix dimensions may be very large, and direct computation of the eigenvalues may be infeasible. One then considers the pencil $A+\mu B$ and attempts to compute a Taylor series expansion to the smallest eigenvalue $\lambda_1(\mu)$ of $A+\mu B$. Generically, for real μ there are no double eigenvalues. The Taylor series reads

$$\lambda_1(\mu)=\sum_{k=0}^{\infty}c_k\mu^k, \tag{5.2}$$

and one then tries to evaluate this at $\mu=\mu_{\mathrm{phys}}$. It is easy to show that $c_0=\lambda_1(0)$, which is the smallest eigenvalue of A (assumed to be simple), and that $c_1=u^*Bu$, with u being a normalized eigenvector of A corresponding to $\lambda_1(0)$. The expressions



(a) $|\mu - \mu(\varepsilon)|$ (b) $|\lambda - \lambda(\varepsilon)|$

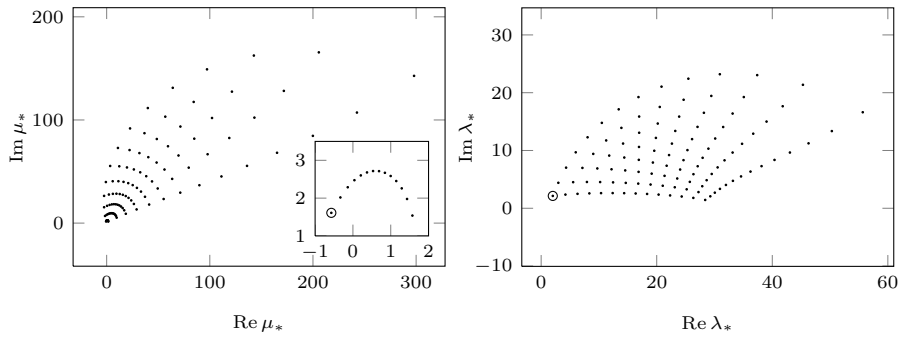
FIGURE 5.1. *Logarithmic plot of the acuracy in the MFRD approach as a function of ε for the example in Section 5.1. The figures show the error in μ_* (left) and λ_* (right), respectively. The lines with markers are the errors for the semisimple point (μ_0, λ_0), while the other lines are the non-semisimple points. In the μ_* and λ_* plot, the vertical lines are at $\varepsilon = \varepsilon_{\text{mach}}^{1/3}$ and $\varepsilon = \varepsilon_{\text{mach}}^{1/2}$, respectively.*

for c_k, $k > 1$ become increasingly complicated, and are typically evaluated using diagrammic rules.

The eigenvalue functions $\lambda_j(\mu)$ are algebraic functions having only branch point type singularities, which is generic for non-semisimple eigenvalues, at complex values of μ. These may occur only when $\lambda_j(\mu_*) = \lambda_k(\mu_*) = \lambda_*$ for $j \neq k$ [SW73]. Consequently, $\min\{|\mu_*|\}$ gives a lower bound for the (basically unknown) radius of convergence of the series (5.2), and the method of fixed relative distance can be used to assess this since all branch points are approximated.

We visualize the set $\mathbb{B}(A, B)$ in Figure 5.2 for a simple example of dimension $n = 15$. The branch point limiting the radius of convergence is highlighted. Here, $\lambda_* = \lambda_1 = \lambda_2$. The example is that of two electrons in a harmonic oscillator trap with the discretzation described in [WKH04]. The matrix A is diagonal with equally spaced eigenvalues $1, 3, 5, \cdots$, while B is dense with matrix elements that decay algebraically, i.e., $B_{ij} = O[(ij)^{-\beta}]$, $\beta > 0$. The matrices thus have a high degree of structure clearly reflected in Figure 5.2.

6. Conclusions and outlook. An important and unusual property of the method in this paper is the *global convergence* to all solutions, which can be combined with an iterative method to get full accuracy. The global convergence property comes from the observation that the fixed relative distance problem is a two-parameter eigenvalue problem. We conclude this paper with some straightforward extensions following from the connection with the two-parameter eigenvalue problem. The method of fixed relative distance can be adapted to generalized eigenvalue problems. That is, the problem of finding (λ, μ) such that $\lambda Cx = (A + \mu B)x$, $x \in \mathbb{C}^n \backslash \{0\}$, where λ is a double eigenvalue, can also be solved by considering the fixed relative distance and solving a two-parameter eigenvalue problem.



(a) μ_* (b) λ_*

FIGURE 5.2. *The set $\mathbb{B}(A, B)$ for a discretized partial differential operator from quantum mechanics. The left plot shows μ_* with the inset figure showing a magnification of the points closest to 0, while the right plot shows λ_*. The radius of convergence of the perturbation series is limited by the magnitude of the marked point in the μ_* plot, for which $\lambda_1(\mu_*) = \lambda_2(\mu_*) = \lambda_*$, the marked point in the λ_* plot.*

The multiparameter eigenvalue problem is a generalization of the two-parameter eigenvalue problem. Consider the problem of finding $\lambda, \mu_1, \mu_2, \ldots, \mu_k$ such that λ is an eigenvalue of multiplicity $k+1$ of the eigenvalue matrix of $A + B_1\mu_1 + \cdots + B_k\mu_k$. This problem can also be solved with an approach based on fixed relative distance.

7. Acknowledgments. We thank Michiel E. Hochstenbach of T.U. Eindhoven, for pointing us to reference [Cot01].

This article present results of the Belgian Programme on Interuniversity Poles of Attraction, initiated by the Belgian State, Prime Minister's Office for Science, Technology and Culture, the Optimization in Engineering Centre OPTEC of the K.U. Leuven, and the project STRT1-09/33 of the K.U. Leuven Research Foundation.

Appendix A. A technical lemma.

LEMMA A.1. *For all $U, V \in \mathbb{C}^{n\times n}$, where either U or V is invertible, we have*

$$\operatorname{adj}(U \otimes V - V \otimes U) = 0. \tag{A.1}$$

Proof. Because the role of U and V can be interchanged we assume, without loosing generality, that V is nonsingular. We further assume that $n \geq 2$, since the result is trivial for $n = 1$.

First we characterize the null space of

$$U \otimes V - V \otimes U. \tag{A.2}$$

From the regularity of V we have

$$(U \otimes V - V \otimes U)X = 0, \quad X \in \mathbb{C}^{n^2\times 1},$$

if and only if

$$((V^{-1}U) \otimes I - I \otimes (V^{-1}U))X = 0,$$

i.e.,

$$((V^{-1}U) \oplus (-V^{-1}U))X = 0. \tag{A.3}$$

If $V^{-1}U$ has (at least) two different eigenvectors, E_1 and E_2, then it follows by inspection that $X = E_1 \otimes E_1$ and $X = E_2 \otimes E_2$ are independent solutions of (A.3). In the other case, there always exists an eigenvalue λ, an eigenvector E and a generalized eigenvector H satisfying

$$(V^{-1}U)E = \lambda E,\ (V^{-1}U)H = \lambda H + E,$$

and it can be verified that $X = E \otimes E$ and

$$X = E \otimes H + H \otimes E$$

are independent solutions of (A.3). Hence, in both cases we conclude that the null space of (A.2) has dimension two. This implies that there exists a matrix $T \in \mathbb{C}^{n^2 \times n^2}$ such that

$$T(U \otimes V - V \otimes U)T^{-1} = \begin{bmatrix} R_{11} & R_{12} \\ 0 & R_{22} \end{bmatrix},$$

where R_{22} is the two-by-two null matrix. We get

$$\operatorname{adj}(T^{-1})\ \operatorname{adj}(U \otimes V - V \otimes U)\ \operatorname{adj}(T) = \operatorname{adj}\left(\begin{bmatrix} R_{11} & R_{12} \\ 0 & R_{22} \end{bmatrix}\right) = 0.$$

The statement of the lemma follows. □

Appendix B. Condition of the solution of (4.1) **and** (4.2)**.** We first consider the condition of the problem of solving (4.1) (in least squares sense). The condition of this problem is closely related to the properties of the Jacobian matrix of (4.1) in the solution, given by

$$J := \begin{bmatrix} M^2 & 0 & (MB + BM)v_1 & -2Mv_1 \\ 0 & M^2 & (MB + BM)v_2 & -2Mv_2 \\ a_1^* & 0 & 0 & 0 \\ a_2^* & 0 & 0 & 0 \\ 0 & a_1^* & 0 & 0 \\ v_2^T & v_1^* & 0 & 0 \end{bmatrix}, \tag{B.1}$$

where $M = \lambda I - A - B\mu$.

In what follows we addresses two important cases.

Case 1: λ is a double, non-semisimple eigenvalue of $A+B\mu$.

There exists a matrix T such that

$$\tilde{M} := T^{-1}MT = \left[\begin{array}{cc|c} 0 & 1 & \\ 0 & 0 & \\ \hline & & R \end{array}\right],$$

where $R \in \mathbb{C}^{(n-2)\times(n-2)}$ is invertible. By pre-multiplying J with $\operatorname{diag}(T^{-1}, T^{-1}, 1, 1, 1, 1)$ and post-multiplying with $\operatorname{diag}(T, T, 1, 1)$ we obtain

$$\tilde{J} := \begin{bmatrix} \tilde{M}^2 & 0 & (\tilde{M}\tilde{B} + \tilde{B}\tilde{M})\tilde{v}_1 & -2\tilde{M}\tilde{v}_1 \\ 0 & \tilde{M}^2 & (\tilde{M}\tilde{B} + \tilde{B}\tilde{M})\tilde{v}_2 & 0 \\ \tilde{a}_1^* & 0 & 0 & 0 \\ \tilde{a}_2^* & 0 & 0 & 0 \\ 0 & \tilde{a}_1^* & 0 & 0 \\ w_2^T & w_1^* & 0 & 0 \end{bmatrix},$$

where

$$\tilde{B} = T^{-1}BT,\ \tilde{v}_1 = T^{-1}v_1,\ \tilde{a}_1^* = a_1^*T,\ \tilde{a}_2^* = a_2^*T,\ w_1^* = v_1^*T,\ w_2^T = v_2^*T.$$

Note that

$$\tilde{M}^2 = \left[\begin{array}{cc|c} 0 & 0 & \\ 0 & 0 & \\ \hline & & R^2 \end{array}\right]. \tag{B.2}$$

It is easy to see that with a random choice of vectors a_1 and a_2 in the normalization constraints in (4.1), the matrix obtained by taking the first two block columns of $\tilde{J}$ has full column rank with probability one, and, in addition, $\tilde{M}\tilde{v}_1 \neq 0$. Therefore, taking into account the structure of (B.2), the matrix $\tilde{J}$ (and thus J) has full column rank if the 2-by-2 matrix obtained by considering the first two rows of

$$[(\tilde{M}\tilde{B} + \tilde{B}\tilde{M})\tilde{v}_1 \ \ -2\tilde{M}\tilde{v}_1]$$

is invertible. This is the case (with probability one for a random choice of a_1 and a_2) if and only if the element at position (2,1) is nonzero, that is,

$$e_2^T(\tilde{M}\tilde{B} + \tilde{B}\tilde{M})\tilde{v}_1 \neq 0, \tag{B.3}$$

where $e_2 = [0\ 1\ 0 \cdots\ 0]^T$ is a unity vector in $\mathbb{C}^n$. Considering that $e_2^T\tilde{M} = 0$ the condition (B.3) becomes, in terms of the original matrices,

$$U_0^*BV_0 \neq 0, \tag{B.4}$$

where $U_0 := T^{-*}e_2$ is the left null vector of M and $V_0 := Mv_1$ the right null vector of M.

The condition (B.4) can be rephrased as

$$\frac{\partial}{\partial\mu}\det(\lambda I - A - B\mu) = 0,$$

which is a necessary and sufficient condition for the complete regular splitting property of the eigenvalue λ of $A + B\mu$, see [HL99].

Recapitulating the above results, we arrive at the following proposition.

PROPOSITION B.1. *Let $(\lambda, \mu) \in \mathbb{B}(A, B)$ be such that λ is a double, non-semisimple eigenvalue of $A + B\mu$ satisfying the completely regular splitting property. Then the problem of solving the equations* (4.1) *is well conditioned, in the sense that the Jacobian matrix corresponding to (λ, μ) generically has full column rank.*

We should note here that the value of the condition number of the Jacobian depends on the choice of the vectors a_1 and a_2 in the normalization.

Case 2: λ is a double semisimple eigenvalue of $A+B\mu$.

The Jacobian (B.1) cannot be of full column rank in the solution of (4.1) because $Mv_1 = 0$. Thus, the problem of solving (4.1) is ill conditioned. However, using the same arguments as spelled out in the previous case, it can be shown that the Jacobian of (4.2) in the corresponding solution generically has *full* column rank. This indicates that the problem of solving (4.2) is *well conditioned*, and it guarantees that Gauss-Newton converges quadratically.

It is important to remark that the element of $\mathbb{B}(A,B)$ under consideration is highly sensitive with respect to changes of A and B, see Proposition 2.5. This is not in contradiction with the above result. The underlying reason is that the high sensitivity of the element of $\mathbb{B}(A,B)$ is due to perturbations of A and B that destroy the property of the presence of a *semi-simple* double eigenvalue. Elements of $\mathbb{B}(A,B)$ that correspond to a non-semisimple double eigenvalue do not solve (4.2), even not in least squares sense.

REFERENCES

[AB05] R. Alam and S. Bora. On sensitivity of eigenvalues and eigendecompositions of matrices. *Linear Algebra Appl.*, 396:273–301, 2005.

[ARK08] B. Adhikari, A. Rafikul, and D. Kressner. Structured eigenvalue condition numbers and linearizations for matrix polynomials. Technical report, ETH Zürich, 2008.

[Atk68] F. Atkinson. Multiparameter spectral theory. *Bull. Am. Math. Soc.*, 74:1–27, 1968.

[Atk72] F. Atkinson. *Multiparameter Eigenvalue Problems. Volume I: Matrices and compact operators*. Academic press, New York, 1972.

[Bau85] H. Baumgärtel. *Analytic perturbation theory for matrices and operators*. Operator Theory: Advances and applications Vol 15. Birkhäuser Verlag, 1985.

[BLO03] J. V. Burke, A. S. Lewis, and M. Overton. Optimization and pseudospectra, with applications to robust stability. *SIAM J. Matrix Anal. Appl.*, 25(1):80–104, 2003.

[CCHI00] F. Chaitin-Chatelin, A. Harrabi, and A. Ilahi. About Hölder condition numbers and the stratification diagram for defective eigenvalues. *Math. Comput. Simul.*, 54:397–402, 2000.

[Cot01] N. Cottin. Dynamic model updating: A multiparameter eigenvalue problem. *Mechanical Systems and Signal Processing*, 15(4):649–665, 2001.

[GW76] G. Golub and J. Wilkinson. Ill-conditioned eigensystems and the computation of the Jordan canonical form. *SIAM Rev.*, 18:578–619, 1976.

[HJ91] R. Horn and C. Johnson. *Topics in Matrix Analysis*. Cambridge University Press, Cambridge, UK, 1991.

[HKP05] M. E. Hochstenbach, T. Košir, and B. Plestenjak. A Jacobi–Davidson type method for the two-parameter eigenvalue problem. *SIAM J. Matrix Anal. Appl.*, 26(2):477–497, 2005.

[HL99] R. Hryniv and P. Lancaster. On the perturbation of analytic matrix functions. *Integral Equations Oper. Theory*, 34(3):325–338, 1999.

[HP03] M. E. Hochstenbach and B. Plestenjak. Backward error, condition numbers, and pseudospectra for the multiparameter eigenvalue problem. *Linear Algebra Appl.*, 375:63–81, 2003.

[HP05] D. Hinrichsen and A. J. Pritchard. *Mathematical Systems Theory I*. Springer-Verlag Berlin, 2005.

[Kar06] M. Karow. Eigenvalue condition numbers and a formula of Burke, Lewis and Overton. *Electron. J. Linear Algebra*, 15:143–153, 2006.

[Kat95] T. Kato. *Perturbation theory for linear operators*. Springer-Verlag, 1995.

[KPM09] D. Kressner, M. J. Peláez, and J. Moro. Structured Hölder condition numbers for multiple eigenvalues. *SIAM J. Matrix Anal. Appl.*, 31(1):175–201, 2009.

[Mai06] A. A. Mailybaev. Computation of multiple eigenvalues and generalized eigenvectors for matrices dependent on parameters. *Numer. Linear Algebra Appl.*, 13:419–436, 2006.

[Mal99] A. N. Malyshev. A formula for the 2-norm distance from a matrix to the set of matrices with multiple eigenvalues. *Numer. Math.*, 83(3):443–454, 1999.

[MBO97] J. Moro, J. V. Burke, and M. Overton. On the Lidskii-Vishik-Lyusternik perturbation theory for eigenvalues of matrices with arbitrary jordan structure. *SIAM J. Matrix Anal. Appl.*, 18(4):793–817, 1997.

[MGWN06] W. Michiels, K. Green, T. Wagenknecht, and S.-I. Niculescu. Pseudospectra and stability radii for analytic matrix functions with application to time-delay systems. *Lin. Alg. Appl.*, 418(1):315–335., 2006.

[MP09a] A. Muhič and B. Plestenjak. On the quadratic two-parameter eigenvalue problem and its linearization. Technical report, University of Ljubljana, 2009.

[MP09b] A. Muhič and B. Plestenjak. On the singular two-parameter eigenvalue problem. *Electron. J. Linear Algebra*, 18:420–437, 2009.

[Ruh70] A. Ruhe. Properties of a matrix with a very ill-conditioned eigenproblem. *Numer. Math.*, 15:57–60, 1970.

[Sun92] J. Sun. On condition numbers of a nondefective multiple eigenvalue. *Numer. Math.*, 61(2):265–275, 1992.

[SW73] T. H. Schucan and H. A. Weidenmüller. Perturbation theory for the effective interaction in nuclei. *Ann. Physics*, 76:483–501, 1973.

[WKH04] A. Wensauer, M. Korkusinski, and P. Hawrylak. Configuration interaction method for Fock-Darwin states. *Solid State Comm.*, 130:115–120, 2004.