UNBOUNDED EIGENVALUE PROBLEMS IN THE GENERALIZED FORM

MARKO HUHTANEN* AND PAULIINA UUSITALO†

Abstract. Unbounded eigenvalue problems in the generalized form are reformulated, implicitly leading to bounded standard eigenvalue problems. By performing implicit linear fractional transformations, the way this can be done is doubly parametrized by the resolvent set. The reformulated problems are then studied in terms of appropriate quotients which can be used to approximate eigenvalues. Based on the location of these quotients, a case for the notion of symmetric eigenvalue problem is made by inspecting an unbounded quadratic form involving two operators. A notion of self-adjointness then arises through reality of the spectrum. The structure proposed extends the classical self-adjointness of a single operator in a natural way. Orthogonality of eigenvectors takes place together with variational characterizations of eigenvalues. In approximating eigenvalues, trial functions are generated by iterating with a certain type of linear fractional transformation, potentially giving rise to cubic convergence in the self-adjoint case proposed.

Key words. PDE eigenvalue problem, Rayleigh quotient, optimal quotient, quadratic form, self-adjointness, variational characterization, cubic convergence, Sturm-Liouville problem, Kamke problem

AMS subject classifications. 47A75, 49R05, 35P15, 47A12

1. Introduction.

This paper is concerned with the eigenvalue problem

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

where A and B are densely defined linear operators in a complex separable Hilbert space $\mathcal{H}$ such that the intersection $D(A) \cap D(B)$ of their domains is dense. Involving two operators, associate with the problem the operator space

$$\mathcal{V} = \operatorname{span}\{A, B\} \tag{1.2}$$

defining each of its element to have the domain $D(A) \cap D(B)$. Under an appropriate notion of closedness, consider then reformulating the eigenvalue problem as

$$Tx = \lambda Sx, \tag{1.3}$$

where T and S are two suitably chosen, linearly independent elements of $\mathcal{V}$. Being a delicate issue in finite dimensions [23, Chapter 15.3], in the unbounded case the choice is even more intricate. Of course, a reformulation was instrumental in von Neumann's approach to deal with the unbounded self-adjoint standard eigenvalue problem [21]. For another example, taking $T = A$ and $S = \frac{1}{s}(sB - A)$, with s in the resolvent set, yields an implicit Yosida transformation of (1.1). (For the Yosida approximation in semigroup theory, see [24, Chapter 1].) When done aptly, a reformulation turns the original problem into a standard eigenvalue problem for the bounded linear operator

$$Sx \overset{TS^{-1}}{\longmapsto} Tx. \tag{1.4}$$

This does not require applying the inverse which, at any rate, is practically never available. (Application of the inverse was critical in von Neumann's construction [21]

* Faculty of Information Technology and Electrical Engineering, University of Oulu, 90570 Oulu 57, Finland, (Marko.Huhtanen@aalto.fi).

† Faculty of Information Technology and Electrical Engineering, University of Oulu, 90570 Oulu 57, Finland, (Pauliina.Uusitalo@oulu.fi).

of spectral theory for self-adjoint operators in terms of bounded operators; see the comments in [17, p. 389], [34, p. 13] and [11, p. 266].) As a family of implicit linear fractional transformation of A and B, the number of ways this can be done is essentially doubly parametrized by the resolvent set. It can be argued that this implicit reformulation largely determines the properties of the eigenvalue problem (1.3). In practice this takes place by inspecting the action of T and S in terms of the associated Rayleigh quotients

$$\mathrm{rq}_{T,S}(x) = \frac{(Tx, Sx)}{\|Sx\|^2} \tag{1.5}$$

although we are more concerned with the optimal quotients

$$\mathrm{oq}_{T,S}(x) = \frac{(Tx, Sx)}{|(Tx, Sx)|} \frac{\|Tx\|}{\|Sx\|} \tag{1.6}$$

for nonzero vectors $x \in D(A) \cap D(B)$. (Here $(\cdot, \cdot)$ denotes the inner-product on $\mathcal{H}$.) Namely, information provided by optimal quotients always surpasses that of Rayleigh quotients; see Corollary 3.8. This paper is concerned with classifying eigenvalue problems based on the location of these quotients. In particular, they simultaneously suggest a notion for self-adjointness. Involving these quotients, iterative and variational methods for approximating eigenvalues are devised. Then cubic convergence is attained in the self-adjoint case.

The optimal quotients were devised to replace the Rayleigh quotients in the numerical solution of very large (matrix) eigenvalue problems [12], leading also to a notion of Hermitianity [13]. For the Rayleigh quotient approximations in the unbounded case, see, e.g., [15], [32, Chapter 4], [9, Chapter 6.5] [26] and [30, Chapter 6] references therein. Quotients are of major importance in nonlinear eigenvalue problems, too; see [19, Section 8] and [18]. Hermitianity is not just a matter of elegance, typically yielding sharper estimates. It can be expected to lead to notable speed-ups in solving practical problems. For eigenvalue problems in the generalized form, bounded or not, both quotients distinctly point at a common notion of self-adjointness. We say that the eigenvalue problem is symmetric if, for some T and S spanning $\mathcal{V}$, the quadratic form

$$x \longmapsto (Tx, Sx), \quad \text{for} \quad x \in D(A) \cap D(B), \tag{1.7}$$

is real valued. If, additionally, the spectrum of (1.3) belongs to $\mathbb{R}$, we say that the eigenvalue problem is self-adjoint. This differs from the traditional setting where T and S are assumed to be self-adjoint operators with S definite. Then S is used to adjust the inner-product; see, e.g, [5, 6].[1] That is, unlike in traditional approaches to deal with generalized eigenvalue problems, now self-adjointness is determined by a single quadratic form (like in the standard case) without any adjustments of the inner-product. And if one wants to do adjustments, the goal in terms of reality of (1.7) is very clearly set.[2] Then non-negativity of (1.7) also provides a natural notion of positive semi-definiteness.

[1] Let us emphasize that in [5, 6] there are no expressions “self-adjoint generalized eigenvalue problem”. It was coined at a later point (of which we are not certain) and is used in numerical linear algebra at least [23].

[2] In concrete terms, there should exists a bounded positive definite operator P such that $x \longmapsto (PTx, Sx)$ is real valued.

Self-adjointness introduced paves way to variational principles for locating and estimating eigenvalues. It is shown that optimal quotients are well-suited to approximating eigenvalues in a gap. Such problem arise, e.g., in relativistic quantum physics leading to indefinite problems [8]. (If one wants to try to use Rayleigh quotients with indefinite problems, various restrictions must be set up; see [22, p. 374] and [8].) Also an orthogonal function theory for eigenvectors involving distinct eigenvalues is shown to exist. In connection with this, generating trial functions, i.e., finite dimensional subspace of $D(A)\cap D(B)$, is a central issue for practical approximation of eigenvalues. An iterative method based on executing a linear fractional transformation involving quotients as

$$x \longmapsto (T-lS)^{-1}(T+\mathrm{oq}_{T,S}(x)S)x$$

is suggested. Here l is either an optimal or Rayleigh quotient. The method can attain cubic convergence in the self-adjoint case, supporting that this notion based on inspecting the quadratic form (1.7) is appropriate. Then, moreover, by imposing appropriate orthogonality conditions, eigenvalues can be found one by one. Consequently, this allows converting the self-adjoint eigenvalue problem into solving linear systems, a task which, at least in principle, is much simpler. Since accurate inversions are not realistic, approximate methods must be considered for a practical method of generating trial functions.

The paper is organized as follows. In Section 2 the set-up is described. Ways to equivalently formulate the problem are suggested. Section 3 deals with the quotients for the generalized eigenvalue problem. Based on these quotients, in Section 4 a notion of self-adjointness for generalized eigenvalue problems is introduced. In Section 5 methods to generate trial functions are described. Iterative and variational techniques for eigenvalue approximation are devised. Experiments are made.

2. Unbounded eigenvalue problems in the generalized form. In formulating generalized eigenvalue problems (1.1), the assumptions should allow covering a large number of applications.[3] (See [25] for one formulation with a certain family of applications in mind.) Then, assuming the intersection of their domains to be dense, the linear operators A and B should be simultaneously closed in an appropriate way. To this end it appears natural to employ the notion of operator space (1.2).

Definition 2.1. *A two dimensional operator space* (1.2)*, with each of its elements defined to have the domain $D(A)\cap D(B)$ which is dense, is said to be closable if it contains two linearly independent closable elements. It is said to be entirely closable if all its elements are closable.*

For the stability of closability in taking linear combinations, see [16, p.190].

If $D(A)\cap D(B) = \mathcal{H}$, then the closability requirement implies that A and B are bounded operators. In particular, then $\mathcal{V}$ is entirely closable.

The standard eigenvalue problem illustrates the closability assumption well.

Example 1. Regarding the standard eigenvalue problem, suppose A is closed and $B = I$, the identity. Then $\mathcal{V}$ is entirely closable. If $D(A) \neq \mathcal{H}$, then the identity operator, which obviously is closable, is the only element of $\mathcal{V}$ whose closure extends beyond $D(A)$ to $\mathcal{H}$.

It is not rare in applications that the domain of B includes that of A. For instance,

[3] In Kato's classic [16, p. 416], the generalized eigenvalue problem is mentioned somewhat in passing.

A can be a differential operator of higher order than B; see, e.g., [14, 6, 2, 29]. The following relies on standard arguments.

PROPOSITION 2.2. *For a closable operator space* (1.2)*, the set of closed operators is open.*

Proof. Assume that $A - \mu B$ with the domain $D(A) \cap D(B)$ is closed for some $\mu \in \mathbb{C}$. Then B is $A - \mu B$ bounded on $D(A) \cap D(B)$, i.e., there exist $a, b \geq 0$ such that

$$\|Bx\| \leq a\|(A - \mu B)x\| + b\|x\| \tag{2.1}$$

for all $x \in D(A) \cap D(B)$. To see this, take $D(A) \cap D(B)$ equipped with the graph norm $\|x\| + \|(A-\mu B)x\|$ making $D(A) \cap D(B)$ complete since $A - \mu B$ is closed. Then B with the domain $D(A) \cap D(B)$ is closed with respect to the graph norm (and thus bounded) so that (2.1) holds.

We have $\|(A - \mu B)x\| \leq \|(A - \lambda B)x\| + \|(\lambda - \mu)Bx\|$. Combine this inequality with (2.1) to have

$$(1 - a|\lambda - \mu|)\|(A - \mu B)x\| \leq \|(A - \lambda B)x\| + |\lambda - \mu|b\|x\|. \tag{2.2}$$

Suppose now that with $x_n \in D(A) \cap D(B)$ we have Cauchy sequences $x_n \to x$ and $(A - \lambda B)x_n \to y$. Then, by using (2.2), we can conclude that for $|\lambda - \mu| < \frac{1}{a}$ we have a Cauchy sequence $(A - \mu B)x_n \to w$. By the closedness of $A - \mu B$ it follows that $x \in D(A) \cap D(B)$. Thus, $A - \lambda B$ with the domain $D(A) \cap D(B)$ is closed if $|\lambda - \mu| < \frac{1}{a}$. ∎

Adjungating the operators yields always an entirely closable $\mathcal{V}$ as follows. (Of course, finding adjoints is not a trivial matter in general.)

EXAMPLE 2. Suppose the linear operators C and D have the domains such that $D(C) \cap D(D)$ is dense. Let $A = C^*$ and $B = D^*$. Then all the elements of $\mathcal{V}$ are closable.

When dealing with a generalized eigenvalue problem, it appears instrumental to define the spectrum in the extended complex plane $\overline{\mathbb{C}}$.[4] Otherwise the following definition conforms with the standard eigenvalue problem formulation for unbounded operators.

DEFINITION 2.3. *A finite* $\lambda \notin \sigma(A, B)$ *if and only if* $A - \lambda B$ *with the domain* $D(A) \cap D(B)$ *is one-to-one and onto* $\mathcal{H}$ *such that the inverse of* $A - \lambda B$ *is bounded. Moreover,* $\infty \notin \sigma(A, B)$ *if and only if* B *with the domain* $D(A) \cap D(B)$ *is one-to-one and onto* $\mathcal{H}$ *such that the inverse of* B *is bounded.*

The resolvent set $\rho(A, B)$ of the eigenvalue problem (1.1) is the complement of $\sigma(A, B)$ in $\overline{\mathbb{C}}$. To avoid trivialities,

$$\sigma(A, B) \neq \overline{\mathbb{C}} \tag{2.3}$$

is assumed. (Of course, establishing this may not be a trivial matter.)

Unless $\mathcal{V}$ entirely closable, there may exist non-closable elements. They correspond to points of the spectrum.

PROPOSITION 2.4. *If* $A - \lambda B$ *is not closable, then* $\lambda \in \sigma(A, B)$.

Proof. Recall that an operator M is closable if and only if $x_n \in D(M)$, $x_n \to 0$ and $Mx_n \to y$ imply $y = 0$ [16, p.165]. So now there exists $x_n \in D(A) \cap D(B)$ and

[4] This is quite standard. For the notion of extended spectrum in the standard self-adjoint case, see [17, p. 390].

a non-zero y such that $x_n \to 0$ and $(A - \lambda B)x_n \to y$. Thus, if the inverse did exist, this would imply $(A - \lambda B)^{-1}y = 0$, a contradiction. □

Whether or not $\infty \in \rho(A, B)$ is the most notable difference between formulating standard eigenvalue problems for bounded and unbounded operators.

PROPOSITION 2.5. *Assume the eigenvalue problem* (1.1) *is standard, i.e.,* $B = I$. *Then* $\infty \in \rho(A, I)$ *if and only if* A *is a bounded operator.*

There are good reasons to consider a reformulation (1.3) of the original eigenvalue problem (1.1) by appropriately choosing T and S so as to have $\infty \in \rho(T, S)$.

EXAMPLE 3. Let $T = A$ and $S = \frac{1}{s}(sB - A)$ with s in the resolvent set $\rho(A, B)$. Then (1.3) is said to be the Yosida formulation of (1.1).

The following is a classical reformulation of the standard self-adjoint case by von Neumann [21].

EXAMPLE 4. Suppose A is self-adjoint and $B = I$. Then the Cayley transformation corresponds to choosing $T = A - iI$ and $S = A + iI$.

Once the inverse of $S = A + iI$ is applied, the Cayley transformation explicitly converts the problem into a bounded (actually unitary) eigenvalue problem. Whenever $\mathcal{V}$ is closable, the underlying construction can be formulated in general by choosing two closable elements as follows.

THEOREM 2.6. *Assume* (1.2) *is closable with* T *closable and* $\infty \in \rho(T, S)$. *Then* TS^{-1} *is a bounded operator on* $\mathcal{H}$ *and*

$$\sigma(T, S) = \sigma(TS^{-1}, I).$$

Proof. Now TS^{-1} is a linear operator on $\mathcal{H}$. It is closable since T is closable.[5] (This follows from the fact that an operator M is closable if and only if $x_n \in D(M)$, $x_n \to 0$ and $Mx_n \to y$ imply $y = 0$ [16, p.165].) Since TS^{-1} is defined on $\mathcal{H}$, it follows that TS^{-1} is closed and hence bounded. This also implies that there are finite $\mu \in \rho(T, S)$. To see this, we have $S - zT = (I - zTS^{-1})S$. Then, since TS^{-1} is bounded, $I - zTS^{-1}$ invertible for $|z|$ small enough. Therefore $(S - zT)^{-1} = S^{-1}(I - zTS^{-1})^{-1}$.

Suppose $\mu \in \rho(T, S)$. Since $S(T - \mu S)^{-1}$ is bounded and $S(T - \mu S)^{-1} = ((T - \mu S)S^{-1})^{-1} = (TS^{-1} - \mu I)^{-1}$, we may conclude that $\mu \in \rho(TS^{-1}, I)$.

Suppose $\mu \in \sigma(T, S)$. If the range of $T - \mu S$ is not $\mathcal{H}$, then neither is the range of $(T - \mu S)S^{-1} = TS^{-1} - \mu I$ and therefore $\mu \in \sigma(TS^{-1}, I)$. If the range of $T - \mu S$ is $\mathcal{H}$, then either $T - \mu S$ has a nullspace or the (formal) inverse of $T - \mu S$ is unbounded. In the former case $(T - \mu S)x = (TS^{-1} - \mu I)Sx = 0$, so that $\mu \in \sigma(TS^{-1}, I)$. In the latter case it follows that $\mu \in \sigma_{ap}(T, S)$ so that there is a sequence of unit vectors $x_j \in D(A) \cap D(B)$ such that $\|Tx_j - \lambda Sx_j\| \to 0$ as $j \to \infty$. Since $\sigma(T, S)$ does not include $\mathbb{C}$, it is not possible that $Tx_j \to 0$ and $Sx_j \to 0$ as $j \to \infty$. (Proof: suppose $Tx_j \to 0$ and $Sx_j \to 0$ as $j \to \infty$. If $\lambda \notin \sigma(T, S)$ then $(T - \lambda S)^{-1}(T - \lambda S)x_j = x_j$. Since $(T - \lambda S)x_j \to 0$ as $j \to \infty$, this would imply that $(T - \lambda S)^{-1}$ is not bounded, a contradiction.) If $\mu \in \sigma_{ap}(T, S)$ is nonzero, then $Tx_j = \mu Sx_j + \epsilon_j$ where $\|e_j\|$ is arbitrarily small for j large enough. Therefore Tx_j and Sx_j are bounded away from zero and $(TS^{-1} - \mu I)Sx_j \to 0$. Consequently, $\mu \in \sigma(TS^{-1}, I)$. If $0 \in \sigma_{ap}(T, S)$, then $TS^{-1}Sx_j \to 0$ and, by the fact that Sx_j must stay bounded away from zero, we have $0 \in \sigma(TS^{-1}, I)$. □

[5] The order is very important here, i.e., if an operator T is closed and S^{-1} is bounded, then $S^{-1}T$ need not be closed. This order has also been suggested [20] although we find it less appealing.

Consequently, once the eigenvalue problem (1.1) is formulated by choosing a basis of the operator space $\mathcal{V}$ in this manner, then $\sigma(T,S)$ is compact. Although hardly ever explicitly available in practice, the bounded operator TS^{-1} carries a lot of information.[6] (The use of the Cayley transformation to study the self-adjoint problem is a manifestation of this.) Its properties are inherited by the reformulated eigenvalue problem (1.3).

Example 5. It is not rare in applications involving a standard eigenvalue problem that the resolvent operator $(sI-A)^{-1}$ is compact; see [16]. Regarding the eigenvalue problem, this does not quite adequately describe the setting. That is, stated in other words, with the choices $T=I$ and $S=sI-A$ this means that operator TS^{-1} is a compact operator. (Or, if we take $T=\mu I-A$, then TS^{-1} equals the identity plus a compact operator.)

It is not realistic to assume having the inverse of S available. Nevertheless, the operator TS^{-1} is accessible in terms of actions on vectors Sx as

$$Sx \overset{TS^{-1}}{\longmapsto} Tx. \tag{2.4}$$

This is an attractive formulation since information in this form is readily available and in many ways sufficient. First and foremost, it can be used to discretize the respective bounded eigenvalue problem (1.3). In particular, it is preferable over the corresponding map related with the original eigenvalue problem (1.1) since (2.4) is bounded. Regarding the degrees of freedom to choose T and S, it is required that $T=A-\alpha B$ and $S=A-\beta B$ with $\alpha\neq\beta$ such that T is closable and $\beta\in\rho(A,B)$. Hence we are dealing with a family of linear fractional transformations of A and B which is doubly parametrized by the resolvent set; see Proposition 2.4.

Example 6. Suppose A is self-adjoint and $B=I$. As is well-known, von Neumann studied the problem in terms of a bounded unitary operator [21]. Then the choices of T and S in the (implicit) Cayley transform correspond to the linear operator

$$(A+iI)x \longmapsto (A-iI)x \tag{2.5}$$

in (2.4); see [27, p. 357]. This is just a special case of what we regard as the unitary generalized eigenvalue problem; see (4.4).

3. Quotients for the generalized eigenvalue problem. The field of values plays an instrumental role in formulating notions such as self-adjoitness [10, Chapter 22] and closedness [16, Chapter 5.3.2] for a linear operator. See also [3, Chapter 4.6]. Its individual elements, i.e., Rayleigh quotients, are of importance in estimating eigenvalues; see the classical papers [2] and [15]. For discusions, see also [4, pp. 122-125] and [19, pp. 9–10]. In what follows we introduce quotients which can be applied regardless of whether the problem is standard or generalized. Conceptually it is pivotal not to be concerned with a single operator. In the eigenvalue problem (1.1) the essential structure is the two dimensional operator space (1.2). Therefore the quotients must be devised accordingly.

3.1. Rayleigh quotients. As illustrated in the previous section, there are no intrinsic reasons to stick to the original eigenvalue problem (1.1) and therefore let us

[6]To avoid trivialities, it is assumed that T and S are linearly independent.

consider (1.3). In particular, reformulating the original problem is a well-known issue in finite dimensions; see [23, Chapter 15.3] and references therein.

Knowing the action (2.4) is sufficient for recovering the Rayleigh quotient of the linear operator TS^{-1} at Sx to have $\frac{(TS^{-1}Sx,Sx)}{\|Sx\|^2}$, so that we may set

$$\mathrm{rq}_{T,S}(x) = \frac{(Tx, Sx)}{\|Sx\|^2}.$$

Since the application of the inverse of S is entirely avoided, this minor trick gives rise to the following notion for the reformulation (1.3).

Definition 3.1. *Assume* (1.2) *is closable with* T *closable and* $\infty \in \rho(T,S)$. *Then*

$$W(T,S) = \left\{ \frac{(Tx, Sx)}{\|Sx\|^2} \ : \ x \in D(A) \cap D(B) \right\}$$

is the field of values of the eigenvalue problem (1.3).

All the tools for bounded operators become now available. By using Theorem 2.6, we may conclude that $\overline{W(T,S)}$ is a convex set containing $\sigma(T,S)$. For another example, for (2.5) we may apply the following by collecting the facts shown so far.

Proposition 3.2. *Assume* (1.2) *is closable with* T *closable and* $\infty \in \rho(T,S)$. *If* (2.4) *is normal, then* $\overline{W(T,S)}$ *is the convex hull of the spectrum of* (1.3).

It is possible to relax the assumptions in Definition 3.1. If we simply take Sx and use it to compute the inner-products with both sides of (1.3) and impose equality, we obtained the quotients of $W(T,S)$. For this it suffices to assume (1.2) to be closable. Emphasizing the vector Sx in this manner alone appears contrived, though. To overcome this, we proceed as follows.

3.2. Optimal quotients. Besides the Rayleigh quotients just introduced, the following quotients provide useful information. These quotients render, in part, the Rayleigh quotients redundant; see (3.7).

Definition 3.3. *Assume* (1.2) *is closable. Then*

$$\left\{ \frac{(Tx, Sx)}{|(Tx, Sx)|} \frac{\|Tx\|}{\|Sx\|} \mid x \in D(A) \cap D(B) \ \text{ with } \ (Tx, Sx) \neq 0 \right\} \tag{3.1}$$

is said to be the field of optimal quotients of the eigenvalue problem (1.3) *such that if* $Tx = 0$ *(resp.* $Sx = 0$*), define the optimal quotient to have the value* 0 *(resp.* ∞*). The field of optimal quotients is denoted by* $\mathcal{F}(T,S)$.[7]

Individual optimal quotients were devised to provide optimal approximations (in finite but very large dimensions) to eigenvalues with respect to given approximate eigenvectors [12]. They can be argued to supersede the Rayleigh quotients in this task. Expressed operator theoretically[8], the rationale in their derivation is to generate a unit vector $y \in \mathcal{H}$ solving

$$\max_{\|y\|=1} \left(|(\frac{Tx}{\|Tx\|}, y)|^2 + |(\frac{Sx}{\|Sx\|}, y)|^2 \right) \tag{3.2}$$

[7] To avoid trivialities, we assume T and S span $\mathcal{V}$.

[8] The derivation relies on the Hilbert space structure through the inner product.

to have a vector simultaneously best pointing at the directions of Tx and Sx.[9] Observe that at an eigenvector x these vectors are parallel. Interpreted Euclidean geometrically, the solution can be regarded as a bisector of these directions; see Appendix. In terms of this, the respective quotient $\frac{(Tx,y)}{(Sx,y)}$ approximating an eigenvalue then equals

$$\mathrm{oq}_{T,S}(x) = \frac{(Tx,Sx)}{|(Tx,Sx)|}\frac{\|Tx\|}{\|Sx\|}; \tag{3.3}$$

see [12]. In particular, if $x \in D(A) \cap D(B)$ is an eigenvector of the problem (1.3), then this yields the corresponding eigenvalue.

Approximations behave as follows, where the approximate point spectrum of the eigenvalue problem (1.1) is denoted by $\sigma_{ap}(T,S)$.

THEOREM 3.4. *Assume* (1.2) *is closable and* (2.3) *holds. Then*

$$\sigma_{ap}(T,S) \subset \overline{\mathcal{F}(T,S)}$$

and $\sigma_{ap}(T,S) \subset \overline{\left\{ \frac{(Tx,Sx)}{\|Sx\|^2} \;:\; x \in D(A) \cap D(B) \right\}}$.

Proof. If $\lambda \in \sigma_{ap}(T,S)$, then there is a sequence of unit vectors $x_j \in D(A) \cap D(B)$ such that $\|Tx_j - \lambda Sx_j\| \to 0$ as $j \to \infty$. Since $\sigma(T,S)$ does not include $\mathbb{C}$, it is not possible that $Tx_j \to 0$ and $Sx_j \to 0$ as $j \to \infty$. (Proof: suppose $Tx_j \to 0$ and $Sx_j \to 0$ as $j \to \infty$. If $\lambda \notin \sigma(T,S)$ then $(T - \lambda S)^{-1}(T - \lambda S)x_j = x_j$. Since $(T - \lambda S)x_j \to 0$ as $j \to \infty$, this would imply that $(T - \lambda S)^{-1}$ is not bounded, a contradiction.)

If $\lambda \in \sigma_{ap}(T,S)$ is nonzero, then $Tx_j = \lambda Sx_j + \epsilon_j$ where $\|e_j\|$ is arbitrarily small for j large enough. Since Tx_j and Sx_j are bounded away from zero, we have

$$\lim_{j\to\infty} \frac{(Tx_j,Sx_j)}{|(Tx_j,Sx_j)|}\frac{\|Tx_j\|}{\|Sx_j\|} = \lambda.$$

The cases $\lambda = 0$ and $\lambda = \infty$ are handled similarly. ☐

This paves way to variational methods. Assume (1.2) is closable and (2.3) holds. We may then bound the eigenvalue closest to any given point z as

$$\inf_{\lambda \in \mathcal{F}(T,S)} |z - \lambda| \le \inf_{\lambda \in \sigma_{ap}(T,S)} |z - \lambda|.$$

EXAMPLE 7. Suppose T is the forward shift on $l^2(\mathbb{Z})$ and $S = I$. Then we have $\mathcal{F}(T,S) = \sigma(T,S)$.

Consider (2.4). There are good reasons to consider a reformulation (1.3) satisfying the assumption $\infty \in \rho(T,S)$.

THEOREM 3.5. *Assume* (1.2) *is closable. If* $\infty \in \rho(T,S)$, *then*

$$\mathcal{F}(T,S) = \mathcal{F}(TS^{-1}, I).$$

Proof. Make a change of variables as $x = S^{-1}y$ for $y \in \mathcal{H}$, so that

$$\frac{(Tx,Sx)}{|(Tx,Sx)|}\frac{\|Tx\|}{\|Sx\|} = \frac{(TS^{-1}y,y)}{|(TS^{-1}y,y)|}\frac{\|TS^{-1}y\|}{\|y\|}.$$

[9] In finite dimensions the argument partly relies on optimally approximating the generalized Schur decomposition.

Collecting all these fractions yields the claim. □

It might be that $0 \in \rho(T,S)$ instead. This is not a serious issue by the fact that interchanging the roles of T and S gives the reciprocal of (3.3). Therefore

$$\mathcal{F}(S,T) = \frac{1}{\mathcal{F}(T,S)}. \tag{3.4}$$

In particular, in choosing a basis of (1.2), it is recommended that either T or S is taken to be invertible and the other closable.

Example 8. Suppose T is a self-adjoint unbounded invertible operator and $S = I$. Then one should interchange the roles of T and S.

The convexity of the field of values is well-know. This appealing property actually turns into a serious obstacle in producing approximations. (This is well-know in solving indefinite eigenvalue problems [8].) An important feature of the field of optimal quotients is its lack of convexity since it allows locating holes in the (approximate point) spectrum as follows.

Corollary 3.6. *Assume* (1.2) *is closable. If* $\infty \in \rho(T,S)$*, then* $\mathcal{F}(T,S)$ *is included in the origin centered annulus with the inner and outer radii*

$$\inf_{x\in\mathcal{H}} \|TS^{-1}x\| \quad \textit{and} \quad \sup_{x\in\mathcal{H}} \|TS^{-1}x\|.$$

Observe that we have

$$|\mathrm{rq}_{T,S}(x)| \leq |\mathrm{oq}_{T,S}(x)|. \tag{3.5}$$

Next we show that at infinity this becomes equality. Thereby information provided by optimal quotients supersedes that of the Rayleigh quotients. Moreover, then convexity is attained since then the notion coincides with the field of values as follows.

Corollary 3.7. *Assume* (1.2) *is closable with* T *closable and* $\infty \in \rho(T,S)$*. Then*

$$\lim_{\mu\to\infty} \mathcal{F}(T-\mu S,S) + \mu = W(T,S).$$

Proof. First $\mathcal{F}(T-\mu S,S) = \mathcal{F}(TS^{-1}-\mu I,I)$. Thereafter, since TS^{-1} is bounded, we can follow the arguments of [13]. That is, denote TS^{-1} by M and let x be of unit length. By the Pythagorean theorem $\|(M-\mu I)x\|^2$ equals

$$|(Mx,x)-\mu|^2 + \|(M-(Mx,x)I)x\|^2 = |(Mx,x)-\mu|^2 + \|Mx\|^2 - |(Mx,x)|^2.$$

Therefore

$$\frac{((M-\mu I)x,x)}{|((M-\mu I)x,x)|}\|(M-\mu I)x\| =$$

$$\frac{(Mx,x)-\mu}{|(Mx,x)-\mu|}\sqrt{|(Mx,x)-\mu|^2 + \|Mx\|^2 - |(Mx,x)|^2} =$$

$$((Mx,x)-\mu)\sqrt{1+\frac{\|Mx\|^2-|(Mx,x)|^2}{|(Mx,x)-\mu|^2}},$$

so that

$$\frac{((M-\mu I)x,x)}{|((M-\mu I)x,x)|}\|(M-\mu I)x\|+\mu=(Mx,x)+O(\frac{1}{|(Mx,x)-\mu|}) \tag{3.6}$$

holds. ◻

Note that if $\infty \in \rho(T,S)$, then (3.6) yields

$$\lim_{\mu\to\infty} \mathrm{oq}_{T-\mu S,S}(x)+\mu=\mathrm{rq}_{T,S}(x), \tag{3.7}$$

so that the Rayleigh quotients can be recovered from optimal quotients through a limit process.

Corollary 3.8. *Assume A is a closable operator. Then*

$$\lim_{\mu\to\infty} \mathrm{oq}_{A-\mu I,I}(x)+\mu=\mathrm{rq}_{A,I}(x)$$

for any $x\in D(A)$.[10]

In particular, in terms of optimal quotients it is possible to construct approximations to eigenvalues without explicitly applying the inverse of S. How to choose T and S in this is critical.

Theorem 3.9. *Assume $\mathcal{V}$ is entirely closable. If $\infty\in\rho(T,S)$, then*

$$\bigcap_{\mu\in\mathbb{C}} \overline{\mathcal{F}(T-\mu S,S)}+\mu=\sigma_{ap}(T,S). \tag{3.8}$$

Proof. Since $\infty\in\rho(T,S)$, the condition (2.3) is satisfied. Therefore we have $\sigma_{ap}(T-\mu S,S)\subset\overline{\mathcal{F}(T-\mu S,S)}$ by Theorem 3.4 for any μ.

For the converse, suppose $0\notin\sigma_{ap}(T,S)$. We may consider the claim for $\sigma(TS^{-1},I)$ and $\mathcal{F}(TS^{-1},I)$. Let us show $\inf_x \frac{\|TS^{-1}x\|}{\|x\|}$ is strictly positive. (The case $\mu\notin\sigma(T,S)$ is proved similarly using $T-\mu S$ by the fact that $\mathcal{V}$ is entirely closable.) Indeed, if $\inf_x \frac{\|TS^{-1}x\|}{\|x\|}=0$, then there is a sequence of unit vectors $x_j\in\mathcal{H}$ such that $TS^{-1}x_j\to 0$ as $j\to\infty$. But this implies $0\in\sigma_{ap}(T,S)$, a contradiction. ◻

For example, if $T=A-iI$ and $S=A+iI$ with A self-adjoint, then $\sigma_{ap}(T,S)=\sigma(T,S)$ and the spectrum is exactly located with (3.8).

In [13] there are several results on the structure of the field of optimal quotients in the case $\dim\mathcal{H}<\infty$. Most of these constructions extend our unbounded case.

4. Symmetric and self-adjoint generalized eigenvalue problems. The location of the numerical range of an operator plays an instrumental role in defining self-adjointness for standard eigenvalue problems [16, 10, 27]. Regarding eigenvalue problems in the generalized form, one is really dealing with the associated operator space (1.2). That is, starting with (1.1), it may be necessary to inspect its appropriate reformulation (1.3). Once done, consider the quadratic form

$$x\longmapsto(Tx,Sx) \tag{4.1}$$

with $x\in D(A)\cap D(B)$ of unit length.[11] (For unbounded quadratic forms, see [30, Chapter 10] and [28] and references therein.) The image of this quadratic form can

[10] Observe that if A is unbounded, then $\infty\notin\rho(A,I)$. However, it is classical that Rayleigh quotients are still defined and used.

[11] Recall that T and S are assumed to span (1.2).

be argued to inherit the role of numerical range in determining self-adjointness. That is, it determines the argument both for the Rayleigh and optimal quotients.

DEFINITION 4.1. *If the range of the quadratic form* (4.1) *belongs* $\mathbb{R}$*, then the eigenvalue problem* (1.3) *is said to be symmetric.*

By Theorem (3.4), this implies that the approximate point spectrum of the problem is a subset of $\mathbb{R}$.

Replacing T and S with any real (linearly independent) linear combinations of T and S preserves symmetry. Consequently, instead of (1.2), in the symmetric case it is natural to be concerned with

$$\operatorname{span}_{\mathbb{R}}\{T, S\}. \tag{4.2}$$

We then say that this operator space (over $\mathbb{R}$) is symmetric. This conceptual flexibility is very useful in producing approximations to eigenvalues. Regarding numerical solving, it seems evident that such a distinctive property should be preserved in a discretization of the eigenvalue problem.[12] In both cases, the "basis problem" consists of choosing a basis in an optimal way.

EXAMPLE 9. Consider the standard eigenvalue problem, i.e., $B = I$, the identity operator. As is well-know, a necessary and sufficient condition on the symmetry of A reads $(Ax, x) \in \mathbb{R}$ for any $x \in D(A)$. Hence Definition 4.1 extends the classical notion of symmetry to generalized eigenvalue problems without any change of the inner-product used.[13]

PROPOSITION 4.2. *Assume the operator space* (4.2) *is symmetric such that* (2.3) *holds. Then both* $T + iS$ *and* $T - iS$ *are one-to-one.*

Proof. For any $x \in D(A) \cap D(B)$ there holds

$$\|(T \pm iS)x\|^2 = \|Tx\|^2 + \|Sx\|^2 \tag{4.3}$$

by the fact that $(Tx, Sx) \in \mathbb{R}$. Since (2.3) holds, necessarily $\|Tx\|^2 + \|Sx\|^2 \neq 0$ for $x \neq 0$. $\square$

This means that the classical (implicit) Cayley transformation (2.5) generalizes and gives rise to an isometric eigenvalue problem

$$(T - iS)x = \lambda(T + iS)x. \tag{4.4}$$

DEFINITION 4.3. *The eigenvalue problem* (1.3) *is said to be self-adjoint if it is symmetric and* $\sigma(T, S) \subset \mathbb{R}$*.*

See Example 12 below for a self-adjoint Kamke problem. And, of course, there is a huge amount of freedom to tune the inner-product in order to achieve self-adjointness if not intially so.[14] Observe that, until now, tuning the inner-product has been the only option to associate properties of self-adjointness with generalized eigenvalue problems. (For the origins of this, see [5].) An elementary trick is as follows.

EXAMPLE 10. In a Sturm-Liouville eigenvalue problem one deals with the operators

$$T = -\frac{d}{dt}\left(p(t)\frac{d}{dt}\right) + q(t) \quad \text{and} \quad S = w(t),$$

[12] It is noteworthy that this is not the case with FEM discretizations of a standard self-adjoint PDE eigenvalue problem.

[13] As is well-known, there are ways to extend symmetry but the idea relies entirely on changing the inner-product; see, e.g.,[6].

[14] Suppose the eigenvalue problem is know to have real spectrum. It is an interesting problem to find, if possible, an inner-product such that the problem becomes self-adjoint.

where p, q and $w > 0$ are given real valued functions defined on an interval $[a,b] \subset \mathbb{R}$. (For Sturm-Liouville problems, see [1].) Appropriate boundary conditions must also be imposed. In this case, one might want to take the inner-product to be $L^2([a,b]; \frac{1}{w(t)}dt)$ to have a self-adjoint problem (1.3).

If the eigenvalue problem is self-adjoint, then the Cayley transformation (2.5) generalizes and reads

$$U = (T + iS)(T - iS)^{-1} \tag{4.5}$$

by being a unitary operator on $\mathcal{H}$. Involving inversion, the explicit Cayley transformation is rarely available.

With the optimal quotients (1.6), an appealing feature in the self-adjoint case is the disappearance of the fraction involving the inner-product, modulo $\pm$. In particular, it seems natural to call the eigenvalue problem (1.3) positive semi-definite if it is self-adjoint and the quadratic form (4.1) is non-negative. (Respectively, indefinite if the quadratic form attains both negative and positive values.) Then we are left with the fraction of norms

$$x \longmapsto \frac{\|Tx\|}{\|Sx\|} \tag{4.6}$$

to study, where $x \in D(A) \cap D(B)$. This is easier to deal with.

THEOREM 4.4. *Assume* (1.3) *is self-adjoint. Then* $S(T - iS)^{-1}$ *is a bounded normal operator on* $\mathcal{H}$ *and*

$$\inf_x \frac{\|Tx\|}{\|Sx\|} = \sqrt{\frac{1}{\|S(T - iS)^{-1}\|^2} - 1}. \tag{4.7}$$

Proof. We have a unitary $U = (T + iS)(T - iS)^{-1} = (T - iS + 2iS)(T - iS)^{-1}$ and therefore

$$S(T - iS)^{-1} = \frac{1}{2i}(U - I)$$

is a bounded operator. It is clearly normal as well.

Divide both sides in (4.3) by $\|Sx\|^2$ to have $\frac{\|(T+iS)x\|^2}{\|Sx\|^2} = \frac{\|Tx\|^2}{\|Sx\|^2} + 1$. Thereby

$$1 \le \inf_x \frac{\|(T + iS)x\|^2}{\|Sx\|^2} = \inf_y \frac{\|(T + iS)(T - iS)^{-1}y\|^2}{\|S(T - iS)^{-1}y\|^2} = \inf_{\|x\|=1} \frac{1}{\|S(T - iS)^{-1}x\|^2}$$

by using $\|(T + iS)(T - iS)^{-1}y\| = \|y\|$. □

For indefinite problems, non-convexity of $\mathcal{F}(T - sS, S)$ is the key to the following variational principle for locating an eigenvalue nearest to a given point $s \in \mathbb{R}$.

COROLLARY 4.5. *Assume* (1.3) *is self-adjoint with* T *closable and* $\infty \in \rho(T, S)$. *If* $s \in \mathbb{R}$, *then*

$$\inf_x \frac{\|(T - sS)x\|}{\|Sx\|} = \min_{\lambda \in \sigma(T - sS, S)} |\lambda|.$$

Proof. We have

$$S(T - sS - iS)^{-1} = ((T - sS - iS)S^{-1})^{-1} = (TS^{-1} - sI - iI)^{-1}.$$

Now $TS^{-1}-sI-iI$ is invertible. Observe that $(Tx,Sx)=(TS^{-1}y,y)$, with $x=S^{-1}y$, is real. Since $y\in\mathcal{H}$ has no constraints and T is closable, we can conclude $TS^{-1}-sI$ is a bounded self-adjoint operator on $\mathcal{H}$. Consequently, $TS^{-1}-sI-iI$ is normal, so that $\|S(T-sS-iS)^{-1}\|$ equals the reciprocal of the distance of i to the spectrum of $TS^{-1}-sI$. Then use the fact that $\sigma(TS^{-1},I)=\sigma(T,S)$ by Theorem 2.6. Thereby the right hand-side of (4.7) yields the distance. ◻

COROLLARY 4.6. *If* (1.3) *is self-adjoint with* T *closable and* $\infty\in\rho(T,S)$*, then*

$$\bigcap_{s\in\mathbb{R}}\overline{\mathcal{F}(T-sS,S)}+s=\sigma(T,S).$$

Let us illustrate with the standard eigenvalue problem that the choice of quotients is a very delicate and non-trivial matter.

EXAMPLE 11. Suppose A is a self-adjoint unbounded invertible operator and one is interested in estimating its eigenvalues near the origin. Choose $T=I$ and $S=A$, so that appropriate assumptions are in place. Then one is interested in the largest eigenvalues of the eigenvalue problem (1.3). Approximations in terms of the quotients read

$$\mathrm{rq}_{I,A}(x)=\frac{(x,Ax)}{\|Ax\|^2}\quad\text{and}\quad\mathrm{oq}_{I,A}(x)=\frac{(x,Ax)}{|(x,Ax)|}\frac{\|x\|}{\|Ax\|}.$$

If A is positive (or negative) definite, optimal quotients hence give better approximations; see (3.5). For indefinite case the situation is similar, as long as x has some quality. And, by using (3.7), we do have $\lim_{\mu\to\infty}\mathrm{oq}_{I-\mu A,A}(x)+\mu=\mathrm{rq}_{I,A}(x)$.

If we choose $T=A$ and $S=I$, then the situation is more intricate as follows.

PROPOSITION 4.7. *Let A be a closed self-adjoint operator with distinct eigenvalues $\lambda_1<\lambda_2\le\lambda_3\le\lambda_4\le\cdots$. Suppose $\lambda_1<\mathrm{rq}_{A,I}(x)<\lambda_2$. Then either*

$$\lambda_1=\mathrm{oq}_{A-\mu I,I}(x)+\mu$$

for a unique μ with $\mathrm{rq}_{A,I}(x)<\mu<\infty$. Or $\lambda_1<\mathrm{oq}_{A-\mu I,I}(x)+\mu<\mathrm{rq}_{A,I}(x)$ for μ with $\mathrm{rq}_{A,I}(x)<\mu<\infty$. The latter holds always when all the eigenvalues are non-positive.

Proof. The function $\mu\mapsto\mathrm{oq}_{A-\mu I,I}(x)+\mu$ is monotonically decreasing with the limit $\mathrm{rq}_{A,I}(x)$ at infinity. Hence the claimed equality is attained for some μ. Or the inequalities hold.

If the eigenvalues are non-positive, A is bounded. Thereby the absolute value of $\mathrm{oq}_{A-\mu I,I}(x)+\mu$ is bounded by the norm of A, i.e., the absolute value of λ_1. ◻

In practice, not only eigenvalues but also eigenvectors are of interest. In the self-adjoint case orthogonality of the associated eigenvectors takes place as follows.

PROPOSITION 4.8. *Assume* (1.3) *is self-adjoint. If x_1 and x_2 are two eigenvectors associated with different eigenvalues, then*

$$((T-iS)x_1,(T-iS)x_2)=0. \tag{4.8}$$

Proof. With eigenvectors of (1.3) we obtain for (4.4), for $j=1,2$,

$$(T-iS)x_j=\lambda(T+iS)x_j$$

for some $\lambda_1\neq\lambda_2$. Applying (4.5) we can conclude that $(T-iS)x_j$ are eigenvectors of a unitary operator. Since $(T-iS)x_1$ and $(T-iS)x_2$ are associated with differing eigenvalues, they are necessarily orthogonal. ◻

For computations this is of importance as follows. Suppose one eigenvector x_1 has been found. To look for another, any approximate eigenvector x_2 aiming at another eigenvalue should be taken to satisfy the $(T-iS)$-orthogonality condition (4.8). (This can be repeated, i.e., always imposing the $(T-iS)$-orthogonality condition against the eigenvectors found so far allows finding eigenvalues one by one.) In particular, for a self-adjoint eigenvalue problem, the $(T-iS)$-orthogonality condition allows also formulating min-max principles for locating eigenvalues in terms of optimal quotients.[15]

Let us end this section by making some preliminary comments on normality, a notion related with self-adjointness in an obvious way [21]. There are some natural options to suggest but they require operations which do not seem to be readily available. (As opposed to the quadratic form (4.6) which is an extremely simple object to inspect.) One is to inspect (2.4); see [13, Proposition 4.2]. The other is to pursue conditions such as (4.8). Observe though that the notion of unitarity for generalized unbounded eigenvalue problem coincides with Definition 4.3. That is, if (1.3) is such that $\sigma(T,S)$ and $\mathcal{F}(T,S)$ are subsets of the unit circle, then there are linear combinations of T and S which give rise to a self-adjoint eigenvalue problem (and vice versa).

5. Generating trial functions and approximating eigenvalues by iterating. Quotients and variational principles are really not much of practical value if there is no arguable way of generating a reasonable "trial" functions, i.e., finite dimensional subspaces of $D(A) \cap D(B)$. Sometimes one is forced to make an educated guess at this point; see, e.g., [31, p. 147 and Chapter 3.6]. (In practice, for realistic 3D problems in complicated geometries, this is somewhat hopeless.) The reason for this is that the power method, the backbone of numerical methods for eigenvalues, is not applicable[16], so that iterations must rely on entirely different principles. In what follows, an iterative method to generate trial functions is described. Requiring applying the inverse, it cannot be regarded as practical in general. It does, however, suggest ways to do approximations which can be expected to produce useful information superior to educated guessing. Also numerical computations can be devised accordingly since the problem converts exclusively into solving linear equations which, in principle at least, is a simpler task.

5.1. Linear fractional iterations. The method devised in [12] can be concisely expressed in terms of a linear fractional transformation formulated as

$$x \longmapsto (T-lS)^{-1}(T+\mathrm{oq}_{T,S}(x)S)x \tag{5.1}$$

where $x \in D(A) \cap D(B)$. The scalar l is either an optimal quotient

$$\mathrm{oq}_{T-\mu S,S}(x)+\mu \tag{5.2}$$

with $\mu \in \mathbb{C}$ or a Rayleigh quotient (which corresponds to taking $\mu=\infty$).

To analyze this map in the infinite dimensional setting, assume (1.3) is self-adjoint with $\infty \in \rho(T,S)$. Consider the case of l being an optimal quotient with $\mu=0$. Assume having an eigenvector approximation corresponding to a discrete eigenvalue

[15] The quadratic form $x \mapsto ((T-iS)x,(T-iS)x)$ gives rise to an "unbounded" inner-product, i.e., it has the properties of an inner-product but only for $x \in D(A) \cap D(B)$. This unboundedness does not prevent performing orthogonalizations for vectors in $D(A) \cap D(B)$ as usual.

[16] The power method is applicable with bounded operators in case the eigenvalue problem is standard.

$\lambda_1 \neq 0$. Since (2.4) is self-adjoint, we may express the approximate eigenvector as

$$x = S^{-1}(q_1 + \epsilon), \tag{5.3}$$

where $S^{-1}q_1$ is the eigenvector of interest and ϵ is orthogonal to q_1. We have

$$x \longmapsto (T - lS)^{-1}(T + lS)x = S^{-1}(TS^{-1} - lI)^{-1}(TS^{-1} + lI)(q_1 + \epsilon).$$

Taking l to be the basic optimal quotient yields

$$l = \mathrm{oq}_{T,S}(x) = \mathrm{sign}(TS^{-1}(q_1 + \epsilon), SS^{-1}(q_1 + \epsilon))\frac{\|TS^{-1}(q_1 + \epsilon)\|}{\|SS^{-1}(q_1 + \epsilon)\|}$$

$$= \mathrm{sign}(TS^{-1}(q_1 + \epsilon), (q_1 + \epsilon))|\lambda_1| + O(\|\epsilon\|^2).$$

Assuming $\mathrm{sign}(TS^{-1}(q_1 + \epsilon), (q_1 + \epsilon)) = \mathrm{sign}(\lambda_1)$ gives $l = \lambda_1 + O(\|\epsilon\|^2)$ and

$$(TS^{-1} + lI)(q_1 + \epsilon) = 2\lambda_1 q_1 + \hat{\epsilon}$$

with $\hat{\epsilon}$ orthogonal to q_1 and $\|\hat{\epsilon}\| = O(\|\epsilon\|)$. Thus we obtain for (5.1)

$$x \longmapsto S^{-1}(TS^{-1} - lI)^{-1}(TS^{-1} + lI)(q_1 + \epsilon) = S^{-1}(O(\frac{1}{\|\epsilon\|^2})q_1 + \tilde{\epsilon})$$

with $\tilde{\epsilon}$ orthogonal to q_1 and $\|\tilde{\epsilon}\| = O(\|\epsilon\|)$. Consequently, after one iteration with (5.1), the component in the direction of q_1 gets magnified by a factor $O(\frac{1}{\|\epsilon\|^2})$ when written as (5.3). This means that in the self-adjoint case it is possible to attain cubic convergence in approximating eigenvalues with optimal quotients. Such a speed is quite impressive.

Once we have an approximate eigenvector, there are many alternatives to generate approximations with (5.2) by appropriately choosing μ. We know that for large eigenvalues the version involving optimal quotients can be expected to outperform the version involving Rayleigh quotients; see Example 11. For the smallest eigenvalues, when the spectrum is unbounded, the choice is harder as illustrated by Proposition 4.7. That is, the situation is far more delicate than what is the common perception; see, e.g., [15, 26].

The most notable differences in iterating with (5.1) compared with classical approaches [23, 7] are as follows. First, we operate with the inverse on the bisector of the vectors Tx and Sx. Second, the way the quotient is chosen is more delicate and problem dependent. Third, we never use adjoint operators in computations. In particular, in infinite dimensions, the adjoint can be very hard to find.

To formulate some guidelines on choosing l, our suggestions are as follows when looking for an extreme eigenvalue.

- Suppose $\infty \in \rho(T, S)$. Then, without loss of generality, we may assume one is looking for the largest eigenvalue in modulus. In this case choose $l = \mathrm{oq}_{T,S}(x)$.
- Suppose $\infty \in \sigma(T, S)$. Then, without loss of generality, we may assume one is looking for the smallest eigenvalue in modulus. In this case, choose $l = \mathrm{rq}_{T,S}(x)$.

When looking for an eigenvalue inside the spectrum, for example when dealing with an indefinite self-adjoint problem, the issue is more delicate. Based on our not so extensive experience, we suggest using optimal quotients.

Let us illustrate this with an easily repeatable example on finding extreme eigenvalues, to see how just one step can yield huge improvements in approximations.

quotient	$j=0$	$j=1$
$\frac{(Tx_j,Sx_j)}{\|Sx_j\|^2}$	-2.5	-2.467403
$-\frac{\|Tx_j\|}{\|Sx_j\|}$	-2.738612	-2.467416

TABLE 5.1

For Example 12 approximations to $\lambda_1 = -\frac{\pi^2}{4} = -2.467401101\ldots$. *After one step 4-5 correct digits more.*

quotient	$j=0$	$j=1$
$\frac{(Tx_j,Sx_j)}{\|Sx_j\|^2}$	-0.333333	-0.404962
$-\frac{\|Tx_j\|}{\|Sx_j\|}$	-0.365148	-0.405112

TABLE 5.2

For Example 12 approximations to $\hat{\lambda}_1 = -\frac{4}{\pi^2} = -0.4052847344\ldots$. *Initially no correct digits and 3 correct digits after one step.*

EXAMPLE 12. Consider a simple Kamke problem [29, 33] involving the operators

$$A = \frac{d^3}{dt^3} \text{ and } B = \frac{d}{dt} \tag{5.4}$$

with $D(A) \cap D(B) = D(A)$. Here $D(A) \subset L^2(0,1) = \mathcal{H}$ consists of those $x \in W^{3,2}(0,1)$ satisfying the boundary conditions $x(0)+x'(0)=0$, $x'(1)=0$ and $x''(0)=0$. By partial integrating, the eigenvalue problem is found to be symmetric with the choices $T=A$ and $S=B$. Self-adjointness follows by the fact that the eigenvalues are known to be $\lambda_j = -\frac{\pi^2}{4}(2j+1)^2$ for $j=0,1,\ldots$. Aiming at $\lambda_1 = -\frac{\pi^2}{4}$, the smallest eigenvalue in modulus, let us use $x_0(t) = \frac{1}{3}t^3 - t + 1$ as an initial guess. Using $l = \mathrm{rq}_{T,S}(x)$ we obtain the approximations after one step listed in Table 5.1.

Let us now take $T=B$ and $S=A$ instead, so that one is interested in locating $\hat{\lambda}_1 = -\frac{4}{\pi^2}$, now the largest eigenvalue in modulus. Using the same initial guess and $l = \mathrm{oq}_{T,S}(x)$, we obtain the approximations after one step listed in Table 5.2.

Assume (1.3) is self-adjoint and an eigenvector x_1 has been found. For further eigenvalues, the iteration with an approximate eigenvector x_2 should always be taken to satisfy the $(T-iS)$-orthogonality condition (4.8). If λ_1 is simple, then this iteration aims at another eigenvalue.

5.2. Incorporating variational principles. The prescribed scheme can be expected to perform very well even without any subspace constructions, assuming the initial guess x has some quality. In practice, however, applying the inverse in (5.1) is rarely realistic. Relaxing the accuracy of inversion must be somehow compensated, though. A natural option is to build a subspace from the iterates to collect any relevant information. For self-adjoint problems we may then invoke variational principles in terms of Corollary 4.5 to approximate eigenvalues; see Algorithm 1. The guidelines of choosing l at every step are the same as before. The value of s at Step 5 can be fixed or vary at every step to be, for example, l as well. Obviously, there is a lot of freedom in the approximative solving at Step 3, including the option of applying numerical methods. Step 3 being the nontrivial part of the scheme, much effort should be put on this.

In the following example we simplify the operator.

EXAMPLE 13. This continues the previous Kamke example involving the operators (5.4). We take $T=A$ and $S=B$. At Step 3 of Algorithm 1 we simplify the

Algorithm 1 A variational linear fractional iteration for a self-adjoint problem (1.3)

1: for an initial guess x set $\mathbf{V} = \operatorname{span}\{x\}$
2: **while** convergence not attained **do**
3: approximately compute $y = (T - lS)^{-1}(T + \mathrm{oq}_{T,S}(x)S)x$
4: set $\mathbf{V} = \operatorname{span}\{\mathbf{V}, y\}$
5: find x solving $\min_{x \in \mathbf{V}} \frac{\|(T-sS)x\|}{\|Sx\|}$
6: **end while**
7: Use x in computing an optimal or Rayleigh quotient.

quotient	$j=0$	$j=1$	$j=2$	$j=3$
$-\frac{\|Tx_j\|}{\|Sx_j\|}$	-2.7	-2.468	-2.4674014	-2.467401103
$\frac{(Tx_j, Sx_j)}{\|Sx_j\|^2}$	-2.5	-2.46749	-2.46740112	-2.467401104

TABLE 5.3
For Example 13 approximations to $\lambda_1 = -\frac{\pi^2}{4} = -2.467401101\ldots$

inversion operation by dropping the lower order terms so that we always compute $y = T^{-1}(T + \mathrm{oq}_{T,S}(x)S)x = (I + \mathrm{oq}_{T,S}(x)T^{-1}S)x$ instead. This is actually a very good approximation. In Table 5.3 we have listed the optimal and Rayleigh quotients obtained after every iterate. The initial guess was again $x_0(t) = \frac{1}{3}t^3 - t + 1$. The value of s varied by being always the most recent Rayleigh quotient.

If the problem is not self-adjoint, then uncertainties arise. It is clear that the variational Step 5 must be devised differently, though. To measure parallelism, a natural option is to replace x with that of solving

$$\max_{x \in \mathbf{V}} \frac{|(Tx, Sx)|}{\|Tx\|\|Sx\|} \tag{5.5}$$

to find the best approximate eigenvector from the subspace $\mathbf{V}$. Its maximum is one, attained when the vectors become parallel, i.e., when an eigenvector is available. This criterion was successfully used in finite dimensions in [12]. Let us consider an example of this.

EXAMPLE 14. Consider the Kamke problem [25] involving the operators

$$A = \frac{d^3}{dt^3} \text{ and } B = (t^2 - 1)\frac{d}{dt} \tag{5.6}$$

with $D(A) \cap D(B) = D(A)$. Here $D(A) \subset L^2(0,1) = \mathcal{H}$ consists of those $x \in W^{3,2}(0,1)$ satisfying the boundary conditions $x(0) = x'(0) = x''(1) = 0$. With $T = A$ and $S = B$ this is not a self-adjoint problem, so that we replace in Algorithm 1 Step 5 with (5.5). By using $y = T^{-1}(T + \mathrm{oq}_{T,S}(x)S)x = (I + \mathrm{oq}_{T,S}(x)T^{-1}S)x$, we are interested in approximating the smallest eigenvalue $\lambda_1 \approx 5.121669$. Let us use $x_0(t) = t^3 - 3t^2$ as an initial guess. Using $l = \mathrm{rq}_{T,S}(x)$ we obtain the approximations after one step listed in Table 5.4. When $j = 4$, the value of (5.5) was 0.99999988.

Appendix: A notion of bisector of two complex Hilbert space vectors. Assume having two non-orthogonal unit vectors w_1, w_2 of a complex Hilbert space $\mathcal{H}$. Consider

$$\max_{\|y\|=1} \left(|(w_1, y)|^2 + |(w_2, y)|^2\right). \tag{5.7}$$

quotient	$j=0$	$j=1$	$j=2$	$j=3$	$j=4$
$-\frac{\|Tx_j\|}{\|Sx_j\|}$	4.9	5.09	5.1204	5.1209	5.1211
$\frac{(Tx_j,Sx_j)}{\|Sx_j\|^2}$	4.4	5.09	5.1204	5.1209	5.1211

TABLE 5.4
For Example 14 approximations to $\lambda_1 = 5.121669\ldots$

This is solved by

$$y = \frac{1}{\sqrt{2+2|(w_2,w_1)|}}\left(\frac{(w_2,w_1)}{|(w_2,w_1)|}w_1 + w_2\right) \tag{5.8}$$

modulo multiplications by $e^{i\theta}$ with $\theta \in [0, 2\pi)$. For a classical Euclidean geometric interpretation, this can be regarded as the bisector of w_1 and w_2 when regarded as lines, i.e., as elements of the Grassmannian $\mathrm{Gr}_1(\mathcal{H})$. This can be argued by the fact that y belongs to the span of w_1 and w_2 such that

$$|(w_1, y)| = |(w_2, y)|$$

holds.

Conclusions. Unbounded eigenvalue problems in the generalized form have been considered and reformulated, with an eye to covering all conceivable applications. Quotients have been introduced for the purpose of eigenvalue approximation, classification of problems as well as variational and iterative computation of approximate eigenvectors. Optimal quotients provide a flexible tool yielding eigenvalue inclusion regions covering Rayleigh quotients as a special case. Formulated in terms of a single quadratic form, a very natural notion of self-adjointness (equiv. unitarity) arises. Iterative methods are deviced for approximating eigenvalues, attaining cubic convergence for self-adjoint problems if applying inverses is feasible. In lack of accurate inversion, variational methods are incorporated into iterations to get approximations.

REFERENCES

[1] W.O. Amrein, A.M. Hinz, D.B. Pearson (eds), Sturm-Liouville Theory. Past and Present, *Birkhuser-Verlag,* 2005.
[2] N. Aronszajn, *The Rayleigh-Ritz and the Weinstein methods for approximation of eigenvalues. II. Differential operators,* Proc. Nat. Acad. Sci. U.S.A., 34 (1948), pp. 594–601.
[3] J. Blank, P. Exner and M. Havlicek, Hilbert Space Operators in Quantum Physics, *Springer Netherlands,* 2008.
[4] R. Bellman, Introduction to Matrix Analysis, 2nd Ed., *Classics in Appl. Math. 19, SIAM Philadelphia*, 1997.
[5] F. Brauer, *Spectral theory for the differential equation* $Lu = \lambda Mu$*,* Canadian J. Math. 10, (1958), pp. 431–446.
[6] E.A. Coddington and H.S.V. de Snoo, Regular Boundary Value Problems Associated with Pairs of Ordinary Differential Expressions, *Lecture Notes in Mathematics, 858. Springer-Verlag, Berlin-New York,* 1981
[7] S.H. Crandall, *Iterative procedures related to relaxation methods for eigenvalue problems,* Proc. Roy. Soc. London, 207 (1951), pp. 416–423.
[8] M.J. Esteban, M. Lewin and E. Séré, *Variational methods in relativistic quantum mechanics*, Bull. Amer. Math. Soc., 45 (2008), pp. 535–593.
[9] L. Evans, Partial Differential Equations, *Graduate Studies in Mathematics, Vol. 19, Amer. Math. Soc.,* 1998.
[10] P. Halmos, Hilbert space problem book, 2nd Ed., *Springer, New York,* 1982

[11] U. Hashagen, *Die Habilitation von John von Neumann an der Friedrich-Wilhelms- Universität in Berlin: Urteile über einen ungarisch-jüdischen Mathematiker in Deutschland im Jahr 1927,* Historia Mathematica, 37 (2010), pp. 242–280.

[12] M. Huhtanen and V. Kotila, *Optimal quotients for solving large eigenvalue problems,* BIT Numer. Math., 59 (2019), pp. 124–154.

[13] M. Huhtanen and V. Kotila, *Field of optimal quotients and Hermitianity,* Linear Alg. Appl., 563 (2019), pp. 527–547.

[14] E. Kamke, *Über die definiten selbstadjungierten Eigenwertaufgaben bei gewöhnlichen linearen Differentialgleichungen. IV,* Math. Z., 48, (1942), pp. 67–100.

[15] T. Kato, *On the upper and lower bounds of eigenvalues,* J. Phys. Soc. Japan, 4 (1949), pp 334–339.

[16] T. Kato, Perturbation Theory for Linear Operators, *Springer–Verlag, Berlin*, 1980.

[17] P. Lax, Functional Analysis *Wiley-Interscience, New York*, 1996.

[18] P. Lindqvist, *On non-linear Rayleigh quotients,* Potential Anal., 2 (1993), pp. 199–218.

[19] J. Mawhin, *Spectra in mathematics and in Physics: from the dispersion of light to nonlinear eigenvalues,* In CIM bulletin, No. 29, pp. 3–13, Centro Internacional de Mathematica, 2011.

[20] I. Nakić, *On the correspondence between spectra of the operator pencil $A - \lambda B$ and of the operator $B^{-1}A$*, Glasnik Matematiki, 51 (2016), pp. 197–221.

[21] J. von Neumann, *Allgemeine Eigenwerttheorie Hermitescher Funktionaloperatoren,* Math. Annalen, 102 (1929), pp. 49–131.

[22] B. Parlett, *Symmetric matrix pencils,* J. Comput. Appl. Math, 38 (1991), pp. 373–385.

[23] B. Parlett, The Symmetric Eigenvalue Problem, *Classics in Applied Mathematics 20, SIAM, Philadelphia,* 1997.

[24] A. Pazy, Semigroups of Linear Operators and Applications to Partial Differential Equations, *Appl. Math. Sci., 44, Springer-Verlag, New York,* 1983.

[25] W.V. Petryshyn, *On the eigenvalue problem $Tu - \lambda Su = 0$ with unbounded and nonsymmetric operators T and S,* Philos. Trans. Roy. Soc. London Ser. A, 262 1967/1968, pp. 413–458.

[26] M. Plum, *Guaranteed numerical bounds for eigenvalues,* Spectral theory and computational methods of Sturm-Liouville problems (Knoxville, TN, 1996), Lecture Notes in Pure and Appl. Math., vol. 191, Dekker, New York, 1997,

[27] W. Rudin, Functional Analysis, 2nd ed., *McGraw-Hill, New York,* 1991.

[28] V. Serov, Fourier Series, Fourier Transform and Their Applications to Mathematical Physics, *Applied Mathematical Sciences, 197. Springer, Cham,* 2017.

[29] A. Shkalikov and C. Tretter, *Kamke problems. Properties of the eigenfunctions,* Math. Nachr., 170 (1994), pp. 251–275.

[30] B. Simon, *Tosio Kato's work on non-relativistic quantum mechanics: part 1,* Bull. Math. Sci., 8 (2018), pp. 121–232.

[31] A. Szabo and N.S. Ostlund, Modern Quantum Chemistry: Introduction to Advanced Electronic Structure Theory, *Mineola, N.Y. Dover*, 1996.

[32] G. Teschl, Mathematical Methods in Quantum Mechanics: With Applications to Schrödinger Operators, 2nd Ed., *Graduate Studies in Math. Vol 157, AMS,* 2014.

[33] C. Tretter, *Nonselfadjoint spectral problems for linear pencils $N - \lambda P$ of ordinary differential operators with λ-linear boundary conditions: Completeness results,* Integral Equat. Oper. Theory, 26 (1996), pp. 222-248.

[34] E. Zeidler, Nonlinear functional analysis and its applications. II/A. Linear monotone operators, *Springer-Verlag, New York,* 1990.