Math. Nachr. **281,** No. 6, 786 – 802 (2008) / **DOI** 10.1002/mana.200510641

Variational principles for symmetric bilinear forms

Jeffrey Danciger[*1], **Stephan Ramon Garcia**[** 2], and **Mihai Putinar**[***3]

[1] Department of Mathematics, Building 380, Stanford University, Stanford, CA 94305, USA
[2] Department of Mathematics, 610 N. College Ave., Pomona College, Claremont, CA 91711, USA
[3] Department of Mathematics, University of California at Santa Barbara, Santa Barbara, California, 93106-3080, USA

Received 23 May 2005, accepted 11 January 2006
Published online 5 May 2008

Key words Symmetric bilinear form, Friedrichs operator, singular values, compact operator, compressed Toeplitz operator, Courant principle, minimax theorem.
MSC (2000) 30D55, 47A15

Every compact symmetric bilinear form B on a complex Hilbert space produces, via an antilinear representing operator, a real spectrum consisting of a sequence decreasing to zero. We show that the most natural analog of Courant's minimax principle for B detects only the evenly indexed eigenvalues in this spectrum. We explain this phenomenon, analyze the extremal objects, and apply this general framework to the Friedrichs operator of a planar domain and to Toeplitz operators and their compressions.

1 Introduction

Due to physical motivations, spectral theory was originally developed for sesquilinear forms, although symmetric bilinear forms over the complex field, such as Hankel forms, are also natural to consider. It is the aim of the present note to discuss the significance of successive Rayleigh quotients for compact symmetric bilinear forms on a Hilbert space.

The general principle which guides the present article is the following. Let $B(x, y)$ be a bounded, symmetric (meaning that $B(x, y) = B(y, x)$ for all x, y) bilinear form, acting on a separable complex Hilbert space $\mathcal{H}$. The Riesz representation theorem produces an antilinear operator $F : \mathcal{H} \to \mathcal{H}$ which satisfies

$$B(x, y) = \langle x, Fy\rangle$$

for all x, y in $\mathcal{H}$. The square $S = F^2$ is a bounded, positive, linear operator. In the case that $\sqrt{S}$ has discrete spectrum $\sigma_0 \geq \sigma_1 \geq \sigma_2 \geq \cdots \geq 0$, for instance if S is compact, the following variational principle holds:

$$\min_{\operatorname{codim} \mathcal{V}=n} \max_{\substack{x\in\mathcal{V} \\ \| x \|=1}} \operatorname{Re} B(x, x) = \begin{cases} \sigma_{2n} & \text{if} \quad 0 \leq n < \frac{\dim \mathcal{H}}{2}, \\ 0 & \text{otherwise}, \end{cases} \tag{1.1}$$

where the expression $(\dim \mathcal{H})/2$ is to be interpreted as ∞ if $\mathcal{H}$ is infinite-dimensional. We will also identify some of the extremal subspaces $\mathcal{V}$ in the expression above.

The first half of this article is devoted to the necessary background and preliminaries required to prove this and several related variational principles. The second half focuses on two applications to classical problems originating in the works of Friedrichs and Takagi.

Our first application goes back to Friedrichs [6], who derived an important inequality (now known as the Friedrichs inequality) from the spectral theory of the bilinear form

$$B(f, g) = \int_\Omega f(z)g(z)\, dA(z), \tag{1.2}$$

* e-mail: danciger@math.stanford.edu, Phone: +1 818 642 3618, Fax: +1 650 725 4066
** Corresponding author: e-mail: Stephan.Garcia@pomona.edu, Phone: +1 909 607 8648, Fax: +1 909 607 1247
*** e-mail: mputinar@math.ucsb.edu, Phone: +1 805 893 3252, Fax: +1 805 893 2385

where f, g belong to $L_a^2(\Omega, dA)$, the Bergman space of the planar domain Ω (see also [19, 20]).

As a byproduct of our treatment of compact symmetric bilinear forms, we identify (for a wide range of domains Ω) the best constant $c(\Omega) < 1$ and an optimal subspace $\mathcal{V}$ of $L_a^2(\Omega)$ of codimension one for which

$$\left|\int_\Omega f^2\, dA\right| \le c(\Omega) \int_\Omega |f|^2\, dA \tag{1.3}$$

holds for all f in $\mathcal{V}$. It turns out that the optimal constant $c(\Omega)$ coincides with the *second* singular number σ_2 of the bilinear form (1.2). For f in an optimal subspace $\mathcal{V}$ of codimension one, we write $f = u + i\widetilde{u}$, where u and $\widetilde{u}$ are real-valued harmonic functions, and deduce from (1.3) the $L^2(dA)$ bound:

$$\int_\Omega \widetilde{u}^2\, dA \le \frac{1 + c(\Omega)}{1 - c(\Omega)} \int_\Omega u^2\, dA.$$

A second important class of examples is furnished by finite Toeplitz matrices and, more generally, by the compressions of Toeplitz operators (acting on the Hardy space H^2) to coinvariant subspaces for the unilateral shift. In the present note we develop several analytic minimax principles for describing the singular values of such operators.

One of the first authors to recognize the importance of the symmetry of a finite Toeplitz matrix with respect to its second diagonal was Takagi [22], who exploited this symmetry in his study of the Carathéodory–Fejér interpolation problem in complex function theory. For this problem, the related bilinear form we consider is given by

$$B(f, g) = \frac{(ufg)^{(n)}(0)}{n!},$$

where $u(z) = c_0 + c_1 z + \cdots + c_n z^n$ is a prescribed Taylor polynomial at the origin. To be more specific, one derives from Takagi's work and our interpretation that there exists an analytic function F in the unit disk such that

$$F(z) = c_0 + c_1 z + \cdots + c_n z^n + O\big(z^{n+1}\big)$$

and $\| F \|_\infty \le M$ if and only if

$$\max_{\| f \|_2 = 1} \frac{1}{n!} \big|\big(uf^2\big)^{(n)}(0)\big| \le M,$$

where f is a polynomial of degree $\le n$ and $\| f \|_2$ denotes the l^2-norm of its coefficients.

There are many other examples of symmetric bilinear forms that our variational principles apply to and we plan to discuss such ramifications in a separate work. For instance, we do not touch in this paper the general theory of Hankel forms [18], nor the important classes of bilinear forms studied in harmonic analysis [2].

2 Preliminaries and main theorems

Just as one considers the relationship between sesquilinear forms and self-adjoint operators, one can study the relationship between bilinear forms and complex symmetric operators. It is for this reason that we briefly discuss such operators here. This requires several basic definitions.

A *conjugation* on a complex Hilbert space $\mathcal{H}$ is an antilinear operator $C : \mathcal{H} \to \mathcal{H}$ that is *involutive* $\big(C^2 = I\big)$ and *isometric*, meaning that $\langle x, y\rangle = \langle Cy, Cx\rangle$ holds for all x, y in $\mathcal{H}$. If $\mathcal{H}$ is separable then for a fixed C it is not hard to show that there exists an orthonormal basis e_n of $\mathcal{H}$ such that $Ce_n = e_n$ for all n. We refer to such a basis as C-*real* and note that C is simply complex conjugation with respect to this basis since $C(\sum a_n e_n) = \sum \overline{a_n} e_n$. We note here again that all Hilbert spaces considered in this article are assumed to be separable.

We say that a bounded operator $T : \mathcal{H} \to \mathcal{H}$ is C-*symmetric* if $T = CT^*C$ and *complex symmetric* if it is C-symmetric with respect to some conjugation C. These definitions are motivated by the fact that the matrix representation of a C-symmetric operator with respect to a C-real orthonormal basis is symmetric [9]. In particular, an $n \times n$ matrix T is symmetric if and only if $T = CT^*C$ where C denotes the standard conjugation

$$C(z_1, z_2, \ldots, z_n) = (\overline{z_1}, \overline{z_2}, \ldots, \overline{z_n}) \tag{2.1}$$

on $\mathbb{C}^n$ and T^* denotes the adjoint of T with respect to the usual sesquilinear inner product on $\mathbb{C}^n$. Thus complex symmetric operators generalize the notion of complex symmetric matrices.

The class of complex symmetric operators is surprisingly large. It includes all normal operators, Hankel operators, compressed Toeplitz operators (including the compressed shift), and many standard integral operators such as the Volterra operator (see [9, 10] for more examples). In the unbounded context, somewhat confusingly, C-symmetric operators are sometimes referred to as *J-selfadjoint*, although this should not be confused with the notion of J-selfadjointness arising in the theory of Krein spaces.

When dealing with C-symmetric operators, the *bilinear* form

$$[x, y] = \langle x, Cy \rangle \tag{2.2}$$

induced by C is almost as important as the standard sesquilinear form $\langle \cdot\,, \cdot \rangle$. With respect to the new form $[\cdot\,, \cdot]$, C-symmetric operators somewhat resemble selfadjoint operators. For instance, T is C-symmetric if and only if $[Tx, y] = [x, Ty]$ for all x, y in $\mathcal{H}$. As another example, it is not hard to show that the eigenvectors of a C-symmetric operator corresponding to distinct eigenvalues are orthogonal with respect to $[\cdot\,, \cdot]$, although they are not necessarily orthogonal with respect to the original sesquilinear form $\langle \cdot\,, \cdot \rangle$.

Despite the fact that one can make certain analogies between complex symmetric and selfadjoint operators, the relationship is often only superficial. The diverse list of complex symmetric operators given above, along with the fact that a complex symmetric matrix can have any possible Jordan structure (see [7, 12]), indicates that such analogies cannot be carried too far.

Our present interest in complex symmetric operators stems from the fact that, for a fixed conjugation C, there is a bijective correspondence between bounded, symmetric bilinear forms $B(x, y)$ on $\mathcal{H} \times \mathcal{H}$ and bounded C-symmetric operators:

Lemma 2.1 *If $B : \mathcal{H} \times \mathcal{H} \to \mathbb{C}$ is a bounded, bilinear form and C is any conjugation on $\mathcal{H}$, then there exists a unique bounded operator T on $\mathcal{H}$ such that*

$$B(x, y) = [Tx, y], \tag{2.3}$$

for all x, y in $\mathcal{H}$, where $[\cdot\,, \cdot]$ denotes the bilinear form (2.2). *If B is symmetric, then T is C-symmetric. Conversely, a bounded C-symmetric operator T gives rise to a bounded, symmetric bilinear form via* (2.3).

Proof. If B is a bounded, bilinear form, then $(x, y) \mapsto B(x, Cy)$ defines a bounded, sesquilinear form on $\mathcal{H} \times \mathcal{H}$ and hence there exists a bounded linear operator $T : \mathcal{H} \to \mathcal{H}$ such that $B(x, Cy) = \langle Tx, y \rangle$ for all x, y in $\mathcal{H}$. Replacing y with Cy in the preceding, we obtain $B(x, y) = [Tx, y]$. If $B(x, y) = B(y, x)$, it follows that $\langle Ty, Cx \rangle = \langle Tx, Cy \rangle$ and hence $\langle x, CTy \rangle = \langle x, T^*Cy \rangle$ holds for all x, y, which shows that T is C-symmetric. Conversely, if T is C-symmetric, then

$$[Tx, y] = \langle Tx, Cy \rangle = \langle x, T^*Cy \rangle = \langle x, CTy \rangle = [x, Ty].$$

The isometric property of C and the Cauchy–Schwarz–Buniakowsky inequality show that the bilinear form $[Tx, y]$ is bounded whenever T is. □

One can also see that the positive operator $|T|$ is uniquely determined by the form B. Indeed, since $B(x, y) = \langle x, T^*Cy \rangle = \langle y, CTx \rangle$, the antilinear operators CT and T^*C are intrinsic to B and thus so is the positive operator

$$(T^*C)(CT) = T^*T = |T|^2.$$

Without any ambiguity, we may therefore make the following definition.

Definition 2.2 A bounded bilinear form $B(x, y)$ is *compact* if the modulus $|T|$ of any of the representing operators T is compact. If $B(x, y)$ is a compact bilinear form, then the *singular values* of B are defined to be the eigenvalues of the positive operator $|T|$, repeated according to their multiplicity.

Of course, if T is any compact operator, then the singular values of T are given by the well-known expression

$$\sigma_n = \min_{\operatorname{codim} \mathcal{V}=n} \max_{\substack{x \in \mathcal{V} \\ \| x \|=1}} \| Tx \|,$$

where $\mathcal{V}$ varies over all $\mathbb{C}$-linear subspaces of $\mathcal{H}$ of codimension n. Considering bilinear forms, however, has advantages in certain situations. For instance, if $\mathcal{H}$ is a Hilbert space of holomorphic functions, then the consideration of bilinear forms often leads to variational characterizations involving explicit holomorphic expressions, rather than a mixture of holomorphic and anti-holomorphic terms (see Sections 3 and 4).

Let us briefly recall Courant's variational principle for the eigenvalues of a self-adjoint operator:

Theorem (Courant) *If A is a compact, self-adjoint operator and $\lambda_0 \geq \lambda_1 \geq \cdots$ are the eigenvalues of A, repeated according to multiplicity, then*

$$\lambda_n = \min_{\operatorname{codim} \mathcal{V}=n} \max_{\substack{x \in \mathcal{V} \\ \| x \|=1}} \langle Ax, x \rangle \tag{2.4}$$

holds whenever $0 \leq n < \dim \mathcal{H}$.

The main result of this note is the following complex symmetric analogue of Courant's principle:

Theorem 2.3 *If T is a compact C-symmetric operator on a separable Hilbert space $\mathcal{H}$ and $\sigma_0 \geq \sigma_1 \geq \cdots \geq 0$ are the singular values of T, repeated according to multiplicity, then*

$$\min_{\operatorname{codim} \mathcal{V}=n} \max_{\substack{x \in \mathcal{V} \\ \| x \|=1}} \operatorname{Re}[Tx, x] = \begin{cases} \sigma_{2n} & \textit{if} \quad 0 \leq n < \frac{\dim \mathcal{H}}{2}, \\ 0 & \textit{otherwise}. \end{cases} \tag{2.5}$$

We defer the proof of Theorem 2.3 until the end of this section. In contrast to Courant's principle, (2.5) reveals only half of the singular values of the given operator. This phenomenon was observed by the first author in the study of the bilinear form $\mathbf{x}^t T \mathbf{y}$ on $\mathbb{C}^n \times \mathbb{C}^n$ induced by a (finite) symmetric matrix T [4]. It turns out that by considering the expression $\operatorname{Re}[Tx, x]$ over $\mathbb{R}$-linear subspaces of $\mathcal{H}$, one avoids this phenomenon and obtains all the singular values of T:

Theorem 2.4 *If T is a compact C-symmetric operator on a separable Hilbert space $\mathcal{H}$ and $\sigma_0 \geq \sigma_1 \geq \cdots \geq 0$ are the singular values of T, repeated according to multiplicity, then*

$$\sigma_n = \min_{\operatorname{codim}_{\mathbb{R}} \mathcal{V}=n} \max_{\substack{x \in \mathcal{V} \\ \| x \|=1}} \operatorname{Re}[Tx, x] \tag{2.6}$$

holds whenever $0 \leq n < \dim \mathcal{H}$. Here $\mathcal{V}$ ranges over all $\mathbb{R}$-linear subspaces of the complex Hilbert space $\mathcal{H}$ and $\operatorname{codim}_{\mathbb{R}} \mathcal{V}$ *denotes the codimension of $\mathcal{V}$ in $\mathcal{H}$ when both are regarded as $\mathbb{R}$-linear spaces.*

In light of Lemma 2.1, we can interpret Theorems 2.3 and 2.4 (proofs of which will be given shortly) in terms of compact, symmetric bilinear forms on $\mathcal{H} \times \mathcal{H}$:

Theorem 2.5 *If $B : \mathcal{H} \times \mathcal{H} \to \mathbb{C}$ is a compact, symmetric bilinear form and if $\sigma_0 \geq \sigma_1 \geq \sigma_2 \geq \cdots \geq 0$ are the singular values of B, repeated according to multiplicity, then*

$$\min_{\operatorname{codim} \mathcal{V}=n} \max_{\substack{x \in \mathcal{V} \\ \| x \|=1}} \operatorname{Re} B(x, x) = \begin{cases} \sigma_{2n} & \textit{if} \quad 0 \leq n < \frac{\dim \mathcal{H}}{2}, \\ 0 & \textit{otherwise}, \end{cases} \tag{2.7}$$

and

$$\min_{\operatorname{codim}_{\mathbb{R}} \mathcal{V}=n} \max_{\substack{x \in \mathcal{V} \\ \| x \|=1}} \operatorname{Re} B(x, x) = \sigma_n \tag{2.8}$$

holds whenever $0 \leq n < \dim \mathcal{H}$.

The final variational principle we consider in this article applies to arbitrary compact (not necessarily symmetric) bilinear forms:

Theorem 2.6 *If $B : \mathcal{H} \times \mathcal{H} \to \mathbb{C}$ is a compact (not necessarily symmetric) bilinear form and if $\sigma_0 \geq \sigma_1 \geq \sigma_2 \geq \cdots \geq 0$ are the singular values of B, repeated according to multiplicity, then*

$$\sigma_n = 2 \min_{\operatorname{codim} \mathcal{V} = n} \max_{\substack{(x,y) \in \mathcal{V} \\ \|(x,y)\| = 1}} \operatorname{Re} B(x, y)$$

for all $n \geq 0$. Here $\mathcal{V}$ denotes a $\mathbb{C}$-linear subspace of $\mathcal{H} \oplus \mathcal{H}$.

Due to the bilinearity of the forms considered, we remark that the expressions $\operatorname{Re} B(x, y)$ and $\operatorname{Re}[Tx, y]$ may be replaced with $|B(x, y)|$ and $|[Tx, y]|$, respectively, in the statements of the theorems above. Furthermore, the proofs of these theorems do not actually require the compactness of T, only the discreteness of the spectrum of $|T|$. It is therefore possible to apply these variational principles if one knows that the spectrum of $|T|$ is discrete. Moreover, it is not hard to see that these variational principles still apply to eigenvalues of $|T|$ located strictly above the essential spectrum of $|T|$.

2.1 Proof of Theorem 2.3

The proof of Theorem 2.3 in the special case where T is a symmetric matrix acting on $\mathbb{C}^n$ (which can be found in [4]) does not immediately extend to the infinite-dimensional setting for several reasons. First, it depends upon a dimension counting argument which requires a minor modification in the infinite dimensional case. Second, it relies heavily upon Takagi's factorization theorem for complex symmetric matrices, which states that a complex symmetric matrix T can be written in the form $T = UDU^t$ where D is the diagonal matrix of singular values of T and U is a unitary matrix (see [12] and the original paper [22]).

Instead of appealing to Takagi's theorem, we require a recent refinement of the polar decomposition for complex symmetric operators due to the second and third authors (see [10]). Recall that the polar decomposition $T = U|T|$ of an operator T expresses T uniquely as the product of a positive operator $|T| = \sqrt{T^*T}$ and a partial isometry U which satisfies $\ker U = \ker |T|$ and maps $\operatorname{cl} \operatorname{ran} |T|$ onto $\operatorname{cl} \operatorname{ran} T$. If T is a C-symmetric operator, then we can decompose the partial isometry U as the product of C with a partial conjugation.

We say that an antilinear operator J is a *partial conjugation* if J restricts to a conjugation operator on $(\ker J)^\perp$ (with values in the same space). In particular, the *linear* operator J^2 is the orthogonal projection onto the closed subspace $\operatorname{ran} J = (\ker J)^\perp$.

The following lemma, whose proof we briefly sketch, is from [10]:

Lemma 2.7 *If $T : \mathcal{H} \to \mathcal{H}$ is a bounded C-symmetric operator, then $T = CJ|T|$ where J is a partial conjugation, supported on $\operatorname{cl} \operatorname{ran} |T|$, which commutes with $|T| = \sqrt{T^*T}$.*

P r o o f. Write the polar decomposition $T = U|T|$ of T and note that

$$T = CT^*C = C|T|U^*C = (CU^*C)(CU|T|U^*C)$$

since U^*U is the orthogonal projection onto $\operatorname{cl} \operatorname{ran} |T|$. One shows that $\ker CU^*C = \ker CU|T|U^*C$, notes that CU^*C is a partial isometry and that $CU|T|U^*C$ is positive, then concludes from the uniqueness of the terms in the polar decomposition that $U = CU^*C$ (so that U is C-symmetric) and that the antilinear operator $J = CU = U^*C$ commutes with $|T|$. One then verifies that J is a partial conjugation supported on $\operatorname{cl} \operatorname{ran} |T|$. □

Without loss of generality, we may assume that J is a conjugation on all of $\mathcal{H}$ since we need only add a partial conjugation supported on the complementary space $\ker |T|$ to obtain a conjugation on all of $\mathcal{H}$ which agrees with J on $\operatorname{cl} \operatorname{ran} |T|$.

Since T is compact, $|T|$ is also compact. The condition $J|T| = |T|J$ implies that each eigenspace of $|T|$ is J-invariant. In particular, J restricts to a conjugation on each of the eigenspaces of $|T|$ and hence (since $J^2 = I$) we can find an orthonormal basis for each eigenspace which is fixed by J. It follows that there exists an orthonormal basis e_n of $\mathcal{H}$ satisfying the conditions

$$|T|e_n = \sigma_n e_n, \quad Je_n = e_n \tag{2.9}$$

for $0 \leq n < \dim \mathcal{H}$.

Step 1: Suppose that $\mathcal{V}$ is a subspace of $\mathcal{H}$ satisfying $\operatorname{codim} \mathcal{V} = n$, where $0 \le n < \frac{\dim \mathcal{H}}{2}$. Considered as a real subspace of $\mathcal{H}$, we have $\operatorname{codim}_{\mathbb{R}} \mathcal{V} = 2n$. Let

$$\mathcal{W} = \operatorname{span}_{\mathbb{R}}\{e_0, e_1, \dots, e_{2n}\}$$

denote the $\mathbb{R}$-linear span of $\{e_0, e_1, \dots, e_{2n}\}$. Write the orthogonal decomposition of the e_j as:

$$e_j = u_j + v_j, \quad u_j \in \mathcal{V}, \quad v_j \in \mathcal{V}^\perp$$

for $j = 0, 1, \dots, 2n$. Here $\mathcal{V}^\perp$ denotes the orthogonal complement of $\mathcal{V}$ with respect to the standard inner product $\langle \cdot, \cdot \rangle$ on $\mathcal{H}$.

Since $\dim_{\mathbb{R}} \mathcal{V}^\perp = 2n$ and the $2n+1$ vectors $v_0, v_1, \dots, v_{2n}$ belong to $\mathcal{V}^\perp$, it follows that there is a nontrivial $\mathbb{R}$-linear combination of them that equals 0:

$$a_0 v_0 + \cdots + a_{2n} v_{2n} = 0.$$

In particular, the coefficients $a_0, a_1, \dots, a_{2n}$ are real and not all zero. Since the vectors $e_0, e_1, \dots, e_{2n}$ are linearly independent (indeed, they are orthonormal) it follows that the vector

$$\begin{aligned} f &= a_0 u_0 + \cdots + a_{2n} u_{2n} \\ &= a_0(e_0 - v_0) + \cdots + a_{2n}(e_{2n} - v_{2n}) \\ &= a_0 e_0 + \cdots + a_{2n} e_{2n} \end{aligned} \tag{2.10}$$

is nonzero, has real coefficients, and belongs to $\mathcal{V} \cap \mathcal{W}$. After normalizing, we may assume that f is a unit vector.

It follows from Lemma 2.7 and Equations (2.9) and (2.10) that

$$\begin{aligned} [Tf, f] &= \langle Tf, Cf \rangle \\ &= \langle CJ|T|f, Cf \rangle \\ &= \langle f, J|T|f \rangle \\ &= \left\langle \sum_{j=0}^{2n} a_j e_j,\, |T| J \sum_{k=0}^{2n} a_k e_k \right\rangle \\ &= \left\langle \sum_{j=0}^{2n} a_j e_j,\, |T| \sum_{k=0}^{2n} \overline{a_k} e_k \right\rangle \\ &= \left\langle \sum_{j=0}^{2n} a_j e_j,\, \sum_{k=0}^{2n} \sigma_k \overline{a_k} e_k \right\rangle \\ &= \sum_{j=0}^{2n} \sigma_j a_j^2 \\ &\ge \sigma_{2n}. \end{aligned}$$

The last line follows from the fact that f is a unit vector and that the a_j are real.

For each subspace $\mathcal{V}$ of $\mathcal{H}$ having (complex) $\operatorname{codim} \mathcal{V} = n$, it therefore follows that

$$\max_{\substack{x \in \mathcal{V} \\ \| x \| = 1}} \operatorname{Re}[Tx, x] \ge \sigma_{2n} \tag{2.11}$$

and hence

$$\min_{\operatorname{codim} \mathcal{V} = n} \max_{\substack{x \in \mathcal{V} \\ \| x \| = 1}} \operatorname{Re}[Tx, x] \ge \sigma_{2n},$$

which concludes the first portion of the proof.

Step 2: For $j = 0, 1, \ldots, n-1$ define the mutually orthogonal vectors

$$w_j = \begin{cases} \sqrt{\sigma_{2j+1}}e_{2j} + i\sqrt{\sigma_{2j}}e_{2j+1} & \text{if} \quad \sigma_{2j+1} > 0, \\ e_{2j+1} & \text{if} \quad \sigma_{2j+1} = 0, \end{cases} \tag{2.12}$$

and consider the $\mathbb{C}$-linear subspace

$$\mathcal{V}' = \operatorname{span}_{\mathbb{C}}\{w_0, w_1, \ldots, w_{n-1}, e_{2n}, e_{2n+1}, \ldots\}, \tag{2.13}$$

which clearly satisfies $\operatorname{codim}_{\mathbb{C}} \mathcal{V}' = n$. Let

$$g = \underbrace{(b_0 w_0 + \cdots + b_{n-1}w_{n-1})}_{g_0} + \underbrace{(c_0 e_{2n} + c_1 e_{2n+1} + \cdots)}_{g_1} \tag{2.14}$$

be a unit vector in $\mathcal{V}'$ and note that

$$Jg = \underbrace{(\overline{b_0}Jw_0 + \cdots + \overline{b_{n-1}}Jw_{n-1})}_{Jg_0} + \underbrace{(\overline{c_0}e_{2n} + \overline{c_1}e_{2n+1} + \cdots)}_{Jg_1}. \tag{2.15}$$

The definition (2.12) of the vectors w_j implies that

$$\langle g_0, |T|Jg_1\rangle = \langle g_1, |T|Jg_0\rangle = 0$$

and hence it follows (using Lemma 2.7 and the computations used in Step 1) that

$$\begin{aligned} [Tg, g] &= \langle g, |T|Jg\rangle \\ &= \langle g_0, |T|Jg_0\rangle + \langle g_0, |T|Jg_1\rangle + \langle g_1, |T|Jg_0\rangle + \langle g_1, |T|Jg_1\rangle \\ &= \langle g_0, |T|Jg_0\rangle + \langle g_1, |T|Jg_1\rangle. \end{aligned} \tag{2.16}$$

Now observe that the term $\langle g_0, |T|Jg_0\rangle$ in (2.16) vanishes altogether. Indeed,

$$\langle w_j, |T|Jw_k\rangle = 0$$

if $j \neq k$ by the definition (2.12) of the w_j and for each $j = 0, 1, \ldots, n-1$ we have

$$\begin{aligned} \langle w_j, |T|Jw_j\rangle &= \left\langle \sqrt{\sigma_{2j+1}}e_{2j} + i\sqrt{\sigma_{2j}}e_{2j+1},\ |T|\big(\sqrt{\sigma_{2j+1}}e_{2j} - i\sqrt{\sigma_{2j}}e_{2j+1}\big)\right\rangle \\ &= \left\langle \sqrt{\sigma_{2j+1}}e_{2j} + i\sqrt{\sigma_{2j}}e_{2j+1},\ \sqrt{\sigma_{2j+1}}\sigma_{2j}e_{2j} - i\sqrt{\sigma_{2j}}\sigma_{2j+1}e_{2j+1}\right\rangle \\ &= \sigma_{2j}\sigma_{2j+1}\langle e_{2j}, e_{2j}\rangle - \sigma_{2j}\sigma_{2j+1}\langle e_{2j+1}, e_{2j+1}\rangle \\ &= 0. \end{aligned}$$

It therefore follows from (2.16) that

$$\begin{aligned} \operatorname{Re}[Tg, g] \leq |[Tg, g]| = |\langle g_1, |T|Jg_1\rangle| &= \left| \sum_{k=0}^{\dim \mathcal{H}-2n} \sigma_{2n+k}c_k^2 \right| \leq \sum_{k=0}^{\dim \mathcal{H}-2n} \sigma_{2n+k}|c_k|^2 \\ &\leq \sigma_{2n} \end{aligned}$$

since g is a unit vector $\big($and hence $\sum |c_k|^2 \leq 1$ by the orthogonality of the terms g_0 and g_1 in (2.14)$\big)$.

For $\mathcal{V}'$ defined by (2.13), it follows from (2.11) that

$$\max_{\substack{x \in \mathcal{V}' \\ \|x\|=1}} \operatorname{Re}[Tx, x] = \sigma_{2n}$$

and hence

$$\min_{\operatorname{codim} \mathcal{V}=n} \max_{\substack{x \in \mathcal{V} \\ \|x\|=1}} \operatorname{Re}[Tx, x] \leq \sigma_{2n},$$

which completes the proof of Theorem 2.3. □

2.2 Proof of Theorem 2.4

The proof of Theorem 2.4 follows the same lines as the proof of Theorem 2.3, and we therefore leave most of the details to the reader. Suppose that $\mathcal{V}$ has real codimension n, where $0 \le n < \dim \mathcal{H}$. Following the comments after Lemma 2.7, we let e_n denote the orthonormal basis of $\mathcal{H}$ satisfying the conditions (2.9). Let $\mathcal{W} = \operatorname{span}_{\mathbb{R}}\{e_0, e_1, \ldots, e_n\}$ denote the $\mathbb{R}$-linear span of the first $n+1$ basis vectors and note that $\mathcal{V}$ and $\mathcal{W}$ intersect nontrivially so that there exists a unit vector $f = a_0e_0 + \cdots + a_ne_n$ in $\mathcal{V}$ with real coefficients $a_0, \ldots, a_n$. This yields $[Tf, f] \ge \sigma_n$ and hence we conclude that

$$\min_{\operatorname{codim}_{\mathbb{R}} \mathcal{V} = n} \max_{\substack{x \in \mathcal{V} \\ \| x \| = 1}} \operatorname{Re}[Tx, x] \ge \sigma_n .$$

One then verifies that equality is achieved for the optimal subspace

$$\mathcal{V}' = \operatorname{span}_{\mathbb{R}}\{ie_0, ie_1, \ldots, ie_n, e_{n+1}, ie_{n+1}, e_{n+2}, ie_{n+2}, \ldots\}.$$

□

2.3 Proof of Theorem 2.6

Fix a conjugation C on $\mathcal{H}$ and let T denote the (not necessarily C-symmetric) compact operator on $\mathcal{H}$ satisfying $B(f, g) = [Tf, g]$. The operator $\widetilde{T}$ on $\mathcal{H} \oplus \mathcal{H}$ is a $\widetilde{C}$-symmetric extension of T, where $\widetilde{T}$ and $\widetilde{C}$ are defined by:

$$\widetilde{T} = \begin{pmatrix} T & 0 \\ 0 & CT^*C \end{pmatrix}, \quad \widetilde{C} = \begin{pmatrix} 0 & C \\ C & 0 \end{pmatrix}.$$

In particular, this shows that every bounded operator on a Hilbert space has a complex symmetric extension.

Since T is compact, it follows that $\widetilde{T}$ is a compact $\widetilde{C}$-symmetric operator. Appealing to Theorem 2.3, we find that if $s_0 \ge s_1 \ge \cdots \ge 0$ denote the singular values of $\widetilde{T}$, then

$$s_{2n} = \min_{\operatorname{codim} \mathcal{V} = n} \max_{\substack{f \in \mathcal{V} \\ \| f \| = 1}} \operatorname{Re} \left\langle \widetilde{T}f, \widetilde{C}f \right\rangle$$

where $\mathcal{V}$ denotes a subspace of $\mathcal{H} \oplus \mathcal{H}$. Letting $f = (x, y)$ denote a typical vector in $\mathcal{H} \oplus \mathcal{H}$, we see that

$$\begin{aligned} \left\langle \widetilde{T}f, \widetilde{C}f \right\rangle &= \langle Tx, Cy\rangle + \langle CT^*Cy, Cx\rangle \\ &= [Tx, y] + \langle x, T^*Cy\rangle \\ &= [Tx, y] + \langle Tx, Cy\rangle \\ &= 2[Tx, y]. \end{aligned}$$

from which it follows that

$$s_{2n} = 2 \min_{\operatorname{codim} \mathcal{V} = n} \max_{\substack{(x,y) \in \mathcal{V} \\ \| (x,y) \| = 1}} \operatorname{Re}[Tx, y]$$

for $n \ge 0$. Since

$$\widetilde{T}^*\widetilde{T} = \begin{pmatrix} T^*T & 0 \\ 0 & CTT^*C \end{pmatrix},$$

it follows that the singular values of $\widetilde{T}$ are simply the eigenvalues of $\sqrt{T^*T}$ and $\sqrt{TT^*}$. Since the nonzero eigenvalues of T^*T and TT^* are the same (the nonzero elements in the spectrum of AB and BA are the same for any operators A and B, see [11, Pr. 76]), it follows that the nonzero eigenvalues of $\left|\widetilde{T}\right|$ are precisely the nonzero eigenvalues of $|T|$, with double the multiplicity. In other words, $s_{2n} = \sigma_n$ whenever $0 \le n < \dim \mathcal{H}$ and hence

$$\sigma_n = 2 \min_{\operatorname{codim} \mathcal{V} = n} \max_{\substack{(x,y) \in \mathcal{V} \\ \| (x,y) \| = 1}} \operatorname{Re}[Tx, y]$$

as claimed.

3 Application: The Friedrichs operator

One instance where symmetric bilinear forms naturally arise without an obvious C-symmetric representing operator is in the consideration of the Friedrichs operator. Although the Friedrichs operator can be pursued in the context of several complex variables [14], we restrict ourselves to the planar case considered in the original article of Friedrichs [6] (this article is also interesting from a historical perspective for it is one of the first in which the Bergman space of a planar domains is explicitly considered). A more modern account of Friedrichs' ideas is contained in the recent articles [15, 19, 20, 21].

Let $\Omega \subset \mathbb{C}$ denote a bounded, connected domain and let $L^2_a(\Omega)$ denote the *Bergman space* of Ω, the Hilbert subspace of all analytic functions in the Lebesgue space $L^2(\Omega) = L^2(\Omega, dA)$. The symmetric bilinear form

$$B(f,g) = \int_\Omega f(z)g(z)\,dA(z) \tag{3.1}$$

on $L^2_a(\Omega) \times L^2_a(\Omega)$ was studied by Friedrichs and others in the context of classical potential theory and planar elasticity. This form is clearly bounded, and it turns out that it is compact whenever the boundary $\partial\Omega$ is $C^{1+\alpha}$ for some $\alpha > 0$ (see [20, 21] for more details). In the other direction, Friedrichs himself showed that if $\partial\Omega$ has an interior angle of α, then $|\sin\alpha/\alpha|$ belongs to the essential spectrum of the form and hence B is not compact. We assume throughout this section that the domain Ω is chosen so that the bilinear form B is compact.

We are interested here in finding the best constant $c(\Omega) < 1$ and an optimal subspace $\mathcal{V}$ of $L^2_a(\Omega)$ of codimension one for which the *Friedrichs inequality*

$$\left|\int_\Omega f^2\,dA\right| \leq c(\Omega)\int_\Omega |f|^2 dA \tag{3.2}$$

holds for all f in $\mathcal{V}$. As we will shortly see, the optimal constant $c(\Omega)$ is precisely σ_2, the second singular value of the bilinear form (3.1).

One important aspect of the Friedrichs inequality is that it provides an $L^2(\Omega, dA)$ bound on harmonic conjugation. Recall that harmonic conjugation $u \mapsto \widetilde{u}$ (where u and $\widetilde{u}$ are real-valued harmonic functions on Ω) is well-defined only after insisting upon a certain normalization for the conjugate functions $\widetilde{u}$. Typically, one requires that $\widetilde{u}$ vanishes at a certain point z_0 in Ω. Such requirements correspond to restricting the analytic function $f = u + i\widetilde{u}$ to lie in a subspace $\mathcal{V}$ of $L^2_a(\Omega)$ of codimension one. The fact that $c(\Omega) = \sigma_2$ in (3.2) yields the best possible $L^2(\Omega, dA)$ bound on harmonic conjugation:

$$\int_\Omega \widetilde{u}^2\,dA \leq \frac{1+\sigma_2}{1-\sigma_2}\int_\Omega u^2\,dA,$$

where $\widetilde{u}$ is normalized so that $u + i\widetilde{u}$ belongs to $\mathcal{V}$. This follows immediately upon substituting $f = u + i\widetilde{u}$ in (3.2) and simplifying (see the proof of Lemma 3.1 for a similar computation).

Without any further restrictions on the domain Ω, the bilinear form (3.1) is not represented by a C-symmetric operator in any obvious way. Indeed, there are few natural conjugations on the Bergman space $L^2_a(\Omega)$ which are evident. Although one might attempt to define a conjugation on $L^2_a(\Omega)$ in terms of complex conjugation with respect to an orthonormal basis of $L^2_a(\Omega)$ $\big($i.e. represent C as the canonical conjugation on $l^2(\mathbb{N})\big)$, such bases are notoriously difficult to describe explicitly, even for relatively simple Ω.

For any fixed conjugation C on $L^2_a(\Omega)$, Lemma 2.1 guarantees the existence of a bounded C-symmetric operator T representing B in the sense that $B(f,g) = [Tf, g] = \langle f, CTg\rangle$ for all f, g in $L^2_a(\Omega)$. In the present situation, it turns out that the antilinear operator CT appearing in the preceding formula is more natural to work with than any potential linear representing operator T.

Let $P_\Omega : L^2(\Omega) \to L^2_a(\Omega)$ denote the *Bergman projection*, the orthogonal projection from the full Lebesgue space $L^2(\Omega)$ onto the Bergman space $L^2_a(\Omega)$. The *Friedrichs operator* is the antilinear operator $F_\Omega : L^2_a(\Omega) \longrightarrow L^2_a(\Omega)$ defined by the equation

$$F_\Omega f = P_\Omega \overline{f},$$

which can also be written in terms of the Bergman kernel $K(z,w)$ of Ω:

$$[F_\Omega f](z) = \int_\Omega K(z,w)\overline{f(w)}\,dA(w), \quad z \in \Omega.$$

The Friedrichs operator represents the bilinear form (3.1) in the sense that

$$B(f,g) = \langle f, F_\Omega g \rangle$$

for all f, g in $L^2_a(\Omega)$. Indeed, this is a straightforward computation:

$$B(f,g) = \langle P_\Omega f, \overline{g} \rangle = \langle f, P_\Omega \overline{g} \rangle = \langle f, F_\Omega g \rangle$$

and hence $CT = F_\Omega$ for any C-symmetric operator T representing the bilinear form B. In light of Lemma 2.7, we see that there exists a canonical partial conjugation J supported on $\operatorname{cl}\operatorname{ran} F_\Omega$ which commutes with $|T|$ and satisfies $F_\Omega = J|T|$.

Since P_Ω is a projection, it follows immediately that $0 \le |T| \le I$. In fact, we can say a good deal more about $|T|$ (or equivalently, about the symmetric bilinear form (3.1)). We start by recalling a useful fact, implicit in the article of Friedrichs:

Lemma 3.1 *If Ω is connected, then $\sigma_1 < \sigma_0 = 1$. In particular, the largest singular value of $B(x,y)$ has multiplicity one and the corresponding eigenfunctions are the constant functions.*

Proof. Since $F_\Omega = J|T|$ and J commutes with $|T|$, one can find a basis of each spectral subspace of $|T|$ (corresponding to a non-zero eigenvalue) which is left invariant by J (see also [10]). If f is such an eigenvector corresponding to the eigenvalue 1, then $|T|f = f$ and $Jf = f$, which implies that $F_\Omega f = f$. Consequently

$$\int_\Omega f^2 \, dA = B(f,f) = \langle f, F_\Omega f \rangle = \langle f, f \rangle = \int_\Omega |f|^2 \, dA.$$

Setting $f = u + iv$ where u and v are real-valued and harmonic, we obtain

$$\int_\Omega \left(u^2 + v^2\right) dA = \int_\Omega \left(u^2 - v^2\right) dA + 2i \int_\Omega uv \, dA = \int_\Omega \left(u^2 - v^2\right) dA$$

since the left-hand side is real. This implies that $\int_\Omega v^2 \, dA = 0$ and hence v vanishes identically on Ω. Since Ω is connected and f analytic, f must be constant throughout Ω. Conversely, it is clear that $\sigma_0 = 1$ since $0 \le |T| \le I$ and F_Ω fixes real constants. □

The following result demonstrates the nature of Friedrichs inequality at the abstract level:

Theorem 3.2 (Abstract Friedrichs Inequality) *If $B : \mathcal{H} \times \mathcal{H} \to \mathcal{H}$ is a compact, symmetric, bilinear form with singular values $\sigma_0 \ge \sigma_1 \ge \cdots \ge 0$, repeated according to multiplicity, and corresponding unit eigenfunctions $e_0, e_1, \ldots$, then*

$$|B(x,x)| \le \sigma_2 \| x \|^2 \tag{3.3}$$

whenever x is orthogonal to the vector $\sqrt{\sigma_1} e_0 + i\sqrt{\sigma_0} e_1$. Furthermore, the constant σ_2 in (3.3) *is the best possible for x restricted to a subspace of $\mathcal{H}$ of codimension one.*

Proof. The result follows from Theorem 2.5 and the construction of the optimal subspace in the proof of Theorem 2.3. □

In essence, (3.3) provides the best possible bound on a symmetric bilinear form that can be obtained on a hyperplane which passes through the origin. We also remark that the orthogonal complement of the vector $\left(\sqrt{\sigma_1} e_0 - i\sqrt{\sigma_0} e_1\right)$ has the same property and hence the optimal subspace in the theorem above is not unique.

The remainder of this section is devoted to several simple illustrations of Theorem 3.2. We restrict our attention to several of the domains for which the associated singular values and eigenfunctions are explicitly computable, relying on computations found in [6] and [19].

Example 3.3 Let Ω_t denote the interior of the ellipse defined by the inequality

$$\frac{x^2}{\cosh^2 t} + \frac{y^2}{\sinh^2 t} < 1, \tag{3.4}$$

where $t > 0$ is a parameter. The basic properties of $\cosh$ and $\sinh$ show that as $t \to \infty$, the domains Ω_t approximate large open disks centered at the origin, while as $t \to 0$, they are the interiors of vertically compressed ellipses, all of which include the interval $[-1, 1]$.

It is well-known that the domains Ω_t satisfy the generalized quadrature identity

$$\int_{\Omega_t} f(z)\, dA(z) = (\sinh 2t) \int_{-1}^{1} f(x)\sqrt{1 - x^2}\, dx$$

(see [3]) and hence the Friedrichs bilinear form $B(f, g) = \langle f, F_{\Omega_t} g\rangle$ can be written as

$$B(f, g) = (\sinh 2t) \int_{-1}^{1} f(x)g(x)\sqrt{1 - x^2}\, dx.$$

In light of this formula, it is not surprising that we encounter Chebyshev polynomials of the second kind. Indeed, the singular values $\sigma_n(t)$ and normalized (in $L^2_a(\Omega_t)$) singular vectors e_n are given by the formulas

$$\sigma_n(t) = \frac{(n + 1)\sinh 2t}{\sinh[2(n + 1)t]},$$
$$e_n(z) = \sqrt{\frac{2n + 2}{\pi \sinh[2(n + 1)t]}}\, U_n(z),$$

where U_n denotes the nth Chebyshev polynomial of the second kind (see [19] or the original paper [6] of Friedrichs).

Since $U_0 = 1$, $U_1 = 2z$, and

$$\sigma_2 = \frac{3\sinh 2t}{\sinh 6t},$$

Theorem 3.2 asserts that if f is orthogonal to the vector

$$\sqrt{\sigma_1}e_0(z) - i\sqrt{\sigma_0}e_1(z) = \frac{2}{\sqrt{\pi \sinh 4t}}(1 - 2iz),$$

then the inequality

$$\left| \int_{\Omega_t} f^2\, dA \right| \le \left(\frac{3\sinh 2t}{\sinh 6t} \right) \int_{\Omega_t} |f|^2\, dA \tag{3.5}$$

holds. Furthermore, the preceding inequality is the best possible that can hold on a subspace of $L^2_a(\Omega_t)$ of codimension one. The relationship between the singular number σ_2 and the parameter t is illustrated in Figure 1.

Since

$$\| 1 - 2iz \|^2_{\Omega_t} = \pi \sinh 2t,$$

the orthogonal projection Q from $L^2_a(\Omega)$ onto the one dimensional subspace spanned by $1 - 2iz$ is readily computed:

$$[Qf](z) = \frac{(1 - 2iz)}{\pi \sinh 2t} \int_{\Omega_t} f(z)(1 + 2i\overline{z})\, dA(z).$$

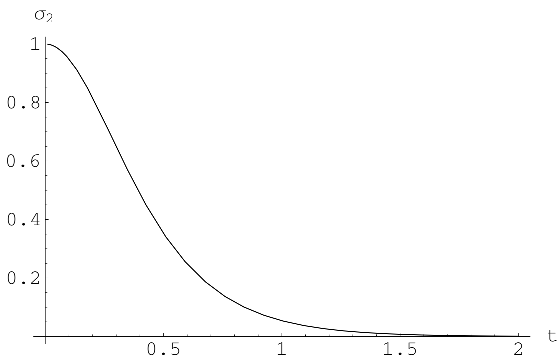

Fig. 1 σ_2 versus t for the ellipses Ω_t (3.4)

Thus, for an arbitrary function f in $L_a^2(\Omega_t)$ we have

$$\left|\int_{\Omega_t} (f - Qf)^2 \, dA\right| \le \left(\frac{3\sinh 2t}{\sinh 6t}\right) \int_{\Omega_t} |f - Qf|^2 \, dA.$$

Similarly, one can replace Qf above by the simpler non-orthogonal projection

$$(Q'f)(z) = \frac{2}{\pi \sinh 2t} \int_{\Omega_t} f(z)(1 + 2i\overline{z}) \, dA(z). \tag{3.6}$$

Indeed, from an arbitrary function f, we wish to construct a function g such that

$$\int_{\Omega_t} g(z)(1 + 2i\overline{z}) \, dA(z) = 0. \tag{3.7}$$

For an arbitrary function f in $L_a^2(\Omega_t)$, we define g by

$$g = f - c \int_{\Omega_t} f(1 + 2i\overline{z}) \, dA$$

where the constant c is chosen so that (3.7) holds. That is, we select c so that

$$\int_{\Omega_t} f(1 + 2i\overline{z}) \, dA - c \left(\int_{\Omega_t} f(1 + 2i\overline{z}) \, dA\right) \int_{\Omega_t} (1 + 2i\overline{z}) \, dA = 0.$$

Since $\int_{\Omega_t} \overline{z} \, dA = 0$, we must choose c so that $c \int_{\Omega_t} \, dA = 1$, yielding (3.6).

Writing $f = u + i\widetilde{u}$, these statements can in turn be interpreted as an $L^2(\Omega_t)$ bound for the harmonic conjugate $\widetilde{u}$ of a square integrable, real harmonic function u on Ω_t.

Example 3.4 Similarly, we can consider the annulus

$$\Omega_t = \left\{z \in \mathbb{C} : e^{-t/2} < |z| < e^{t/2}\right\}, \tag{3.8}$$

where $t > 0$ is a parameter. In this situation, the singular values (apart from σ_0) occur in pairs

$$1 = \sigma_0 > \sigma_1 = \sigma_2 > \sigma_3 = \sigma_4 > \cdots$$

(see [19] for details). In particular,

$$\sigma_1 = \sigma_2 = \sqrt{\frac{\tanh t}{t}}$$

and hence Theorem 3.2 implies that

$$\left| \int_{\Omega_t} f^2 \, dA \right| \;\leq\; \sqrt{\frac{\tanh t}{t}} \int_{\Omega_t} |f|^2 \, dA$$

for all functions f in $L^2_a(\Omega_t)$ which are orthogonal to the constant functions (i.e. the 0th Laurent coefficient of f vanishes). Furthermore, this is the best possible inequality that can hold on a subspace of $L^2_a(\Omega_t)$ of codimension one. The relationship between the singular number σ_2 and the parameter t is illustrated in Figure 2.

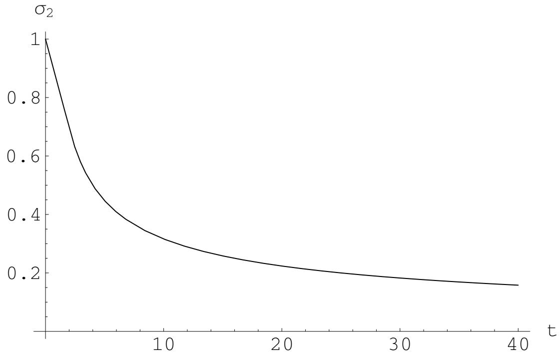

Fig. 2 σ_2 versus t for the annuli Ω_t (3.8)

4 Application: Toeplitz matrices and compressed Toeplitz operators

The variational principles introduced in Section 2 can also be used to characterize singular values (and hence norms) of Toeplitz matrices and, more generally, the compressions of Toeplitz operators to $*$-invariant subspaces of the Hardy space H^2 (the so-called model spaces).

Let T denote a $m \times m$ *Toeplitz matrix*

$$\begin{pmatrix} c_0 & c_{-1} & c_{-2} & \cdots & c_{-(m-1)} \\ c_1 & c_0 & c_{-1} & \cdots & c_{-(m-2)} \\ c_2 & c_1 & c_0 & \cdots & c_{-(m-3)} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ c_{m-1} & c_{m-2} & c_{m-3} & \cdots & c_0 \end{pmatrix} \tag{4.1}$$

regarded as an operator on $\mathbb{C}^m$ with respect to the standard basis. Although T is not in general symmetric with respect to the main diagonal, it is symmetric with respect to the second diagonal. It is not hard to show that this implies that T is C-symmetric with respect to the conjugation

$$C(z_0, z_1, \ldots, z_{m-1}) = (\overline{z_{m-1}}, \overline{z_{m-2}}, \ldots, \overline{z_0}) \tag{4.2}$$

on $\mathbb{C}^m$. Using Theorems 2.3, 2.4, and 2.6, we have the following variational characterization of the singular values of Toeplitz matrices:

Theorem 4.1 *Let T denote the Toeplitz matrix* (4.1) *and let $\sigma_0 \geq \sigma_1 \geq \cdots \geq \sigma_{m-1}$ denote the singular values of T, repeated according to multiplicity. If $\mathbf{z} = (z_0, z_1, \ldots, z_{m-1})$ and $\mathbf{w} = (w_0, w_1, \ldots, w_{m-1})$ denote typical vectors in $\mathbb{C}^m$, then*

$$\sigma_{2n} = \min_{\substack{\mathcal{V} \subset \mathbb{C}^m \\ \operatorname{codim} \mathcal{V} = n}} \max_{\substack{\mathbf{z} \in \mathcal{V} \\ \|\mathbf{z}\|=1}} \operatorname{Re} \sum_{i+j+k=m-1} c_i z_j z_k,$$

$$\sigma_n = \min_{\substack{\mathcal{V} \subset \mathbb{C}^m \\ \operatorname{codim}_{\mathbb{R}} \mathcal{V} = n}} \max_{\substack{\mathbf{z} \in \mathcal{V} \\ \|\mathbf{z}\|=1}} \operatorname{Re} \sum_{i+j+k=m-1} c_i z_j z_k,$$

$$\sigma_n = 2 \min_{\substack{\mathcal{V} \subset \mathbb{C}^m \oplus \mathbb{C}^m \\ \operatorname{codim} \mathcal{V} = n}} \max_{\substack{(\mathbf{z},\mathbf{w}) \in \mathcal{V} \\ \|(\mathbf{z},\mathbf{w})\| = 1}} \operatorname{Re} \sum_{i+j+k=m-1} c_i z_j w_k,$$

where $-(m-1) \leq i \leq m-1$ *as indicated in* (4.1).

Corollary 4.2 *The operator norm of the Toeplitz matrix* (4.1) *is given by*

$$\|T\| = \max_{\substack{\mathbf{z} \in \mathbb{C}^m \\ \|\mathbf{z}\| = 1}} \operatorname{Re} \sum_{i+j+k=m-1} c_i z_j z_k$$

where $-(m-1) \leq i \leq (m-1)$.

These variational characterizations of the norms and singular values of Toeplitz matrices can be considerably generalized. Indeed, Toeplitz matrices are the simplest examples of *compressed Toeplitz operators*, an important class of operators which we briefly discuss below.

Let H^2 denote the Hardy space on the unit disk $\mathbb{D}$, the Hilbert space of all analytic functions on $\mathbb{D}$ with finite norm with respect to the inner product

$$\langle f, g \rangle = \lim_{r \to 1^-} \frac{1}{2\pi} \int_{-\pi}^{\pi} f(re^{it}) \overline{g(re^{it})} \, dt. \tag{4.3}$$

As is well-known, we may freely identify H^2 functions with their boundary functions on the unit circle $\partial\mathbb{D}$ (defined a.e. with respect to Lebesgue measure) and we therefore regard H^2 as the subspace of $L^2 = L^2(\partial\mathbb{D})$ consisting of functions with vanishing negatively indexed Fourier coefficients.

For each u in $L^\infty(\partial\mathbb{D})$, the *Toeplitz operator* with *symbol* u is the operator $T_u : H^2 \to H^2$ defined by $T_u f = P(uf)$, where P denotes the orthogonal projection from L^2 onto H^2. Roughly speaking, applying P to uf truncates the full Fourier series of uf and returns only the nonnegatively indexed terms.

An immediate corollary of Beurling's theorem characterizes the proper, nontrivial invariant subspaces of the backward shift operator on H^2. They are all of the form $H^2 \ominus \varphi H^2$, where φ is a nonconstant *inner function* (a bounded analytic function on the unit disk $\mathbb{D}$ with nontangential limiting values of unit modulus a.e. on $\partial\mathbb{D}$) and φH^2 denotes the subspace $\{\varphi f : f \in H^2\}$. Such spaces are commonly refereed to as $*$-invariant subspaces or *model spaces* (in light of a famous theorem of Sz.-Nagy and Foaiş). See the classics [5, 13] and the recent text [16] for background.

A *compressed Toeplitz operator* is an operator of the form $P_\varphi T_u P_\varphi$ where T_u is a Toeplitz operator and P_φ denotes the orthogonal projection from H^2 onto $H^2 \ominus \varphi H^2$. With a slight abuse of notation, we regard compressed Toeplitz operators as operators acting on $H^2 \ominus \varphi H^2$, rather than on H^2 itself.

The following lemma shows that each model space carries a natural conjugation operator:

Lemma 4.3 *If* φ *is a nonconstant inner function, then*

$$Cf = \overline{fz}\varphi \tag{4.4}$$

defines a conjugation operator on $H^2 \ominus \varphi H^2$.

The proof, which amounts to verifying that $\langle Cf, \overline{zh} \rangle = \langle Cf, \varphi h \rangle = 0$ for all H^2 functions h, can be found in [9] or the expository note [8]. Although the conjugation (4.4) is defined in terms of boundary functions, in the case where φ is a finite Blaschke product, the conjugation can be explicitly described.

Example 4.4 Let φ denote a finite Blaschke product

$$\varphi(z) = \prod_{j=1}^{m} \frac{z - \lambda_j}{1 - \overline{\lambda_j} z} \tag{4.5}$$

with m (not necessarily distinct) zeros λ_j inside of $\mathbb{D}$. A partial fractions argument shows that each function f in $H^2 \ominus \varphi H^2$ has a unique representation in the form

$$f(z) = \frac{a_0 + a_1 z + \cdots + a_{m-1} z^{m-1}}{(1 - \overline{\lambda_1} z) \cdots (1 - \overline{\lambda_m} z)}.$$

A short computation using the definition (4.4) shows that

$$[Cf](z) = \frac{\overline{a_{m-1}} + \overline{a_{m-2}}z + \cdots + \overline{a_0}z^{m-1}}{(1-\overline{\lambda_1}z)\cdots(1-\overline{\lambda_m}z)}.$$

In particular, if $\varphi = z^m$, then $H^2 \ominus z^m H^2$ is simply the subspace $\mathcal{P}_{m-1}$ of H^2 consisting of polynomials of degree $\leq m-1$. The corresponding conjugation is given by

$$C\big(a_0 + a_1 z + \cdots + a_{m-1}z^{m-1}\big) = \overline{a_{m-1}} + \overline{a_{m-2}}z + \cdots + \overline{a_0}z^{m-1}. \tag{4.6}$$

In particular, there is a natural identification of $\mathcal{P}_{m-1} = H^2 \ominus z^m H^2$ with $\mathbb{C}^m$ so that the conjugations (4.2) and (4.6) coincide.

The main result we require now is quoted from [9] (a more detailed analysis of this result and its applications can be found in [8, 10]):

Lemma 4.5 *If φ is a nonconstant inner function, u belongs to $L^\infty(\partial\mathbb{D})$, and P_φ denotes the orthogonal projection from H^2 onto $H^2 \ominus \varphi H^2$, then the compression $P_\varphi T_u P_\varphi$ of the Toeplitz operator T_u to $H^2 \ominus \varphi H^2$ is C-symmetric with respect to the conjugation operator* (4.4).

Although a nonzero Toeplitz operator is never compact [11, Cor. 2, Pr. 242], its compressions to model spaces are often compact. Indeed, if $H^2 \ominus \varphi H^2$ is finite-dimensional (i.e. φ is a finite Blaschke product), then the compression of any Toeplitz operator to $H^2 \ominus \varphi H^2$ is trivially compact. Moreover, if φ is an interpolating Blaschke product with zeroes λ_j, u belongs to H^∞, and $\sum_{j=1}^\infty |u(\lambda_j)|^2 < \infty$, then it follows from Carleson's interpolation theorem that the compression of T_u to $H^2 \ominus \varphi H^2$ is compact. This follows immediately from the fact that the eigenvectors of $P_\varphi T_u P_\varphi$ form a Riesz basis for $H^2 \ominus \varphi H^2$.

We can interpret finite Toeplitz matrices in the present context and obtain an analytic characterization of norms and singular values. The constants c_j in the Toeplitz matrix (4.1) can be interpreted as the Laurent coefficients at 0 of the rational function

$$u(z) = \frac{c_{-(m-1)}}{z^{m-1}} + \cdots + \frac{c_{-1}}{z} + c_0 + c_1 z + \cdots + c_{m-1}z^{m-1} \tag{4.7}$$

and the Toeplitz matrix T can be identified with the compression of the Toeplitz operator T_u to the subspace $\mathcal{P}_{m-1} = H^2 \ominus z^m H^2$ of polynomials of degree $\leq m-1$. After identifying polynomials

$$f(z) = a_0 + a_1 z + \cdots + a_{m-1}z^{m-1}$$

of degree $\leq m-1$ with their coefficient sequences $(a_0, a_1, \ldots, a_{m-1})$ in $\mathbb{C}^m$ it follows that

$$\begin{aligned}[][Tf,g] &= \langle P_\varphi T_u P_\varphi f, Cg\rangle = \langle T_u f, P_\varphi Cg\rangle = \langle P(uf), Cg\rangle = \langle uf, Cg\rangle = \langle uf, \overline{gz}z^m\rangle \\ &= \frac{1}{2\pi i}\int_{\partial\mathbb{D}} \frac{u(\zeta)f(\zeta)g(\zeta)}{\zeta^m}\, d\zeta = \operatorname{Res}_0 \frac{ufg}{z^m}\end{aligned}$$

where Res_0 returns the residue of a function at the origin. This yields the following result:

Theorem 4.6 *If T denotes the $m \times m$ Toeplitz matrix* (4.1)*, $\sigma_0 \geq \sigma_1 \geq \cdots \geq \sigma_{m-1}$ denotes the singular values of T, repeated according to multiplicity, and u denotes the corresponding rational function* (4.7)*, then*

$$\sigma_{2m} = \min_{\substack{\mathcal{V}\subset\mathcal{P}_{m-1}\\ \operatorname{codim}\mathcal{V}=m}} \max_{\substack{f\in\mathcal{V}\\ \|f\|_2=1}} \operatorname{Re}\left(\operatorname{Res}_0 \frac{uf^2}{z^m}\right),$$

$$\sigma_{m} = \min_{\substack{\mathcal{V}\subset\mathcal{P}_{m-1}\\ \operatorname{codim}_{\mathbb{R}}\mathcal{V}=m}} \max_{\substack{f\in\mathcal{V}\\ \|f\|_2=1}} \operatorname{Re}\left(\operatorname{Res}_0 \frac{uf^2}{z^m}\right),$$

$$\sigma_{m} = \min_{\substack{\mathcal{V}\subset\mathcal{P}_{m-1}\oplus\mathcal{P}_{m-1}\\ \operatorname{codim}\mathcal{V}=m}} \max_{\substack{(f,g)\in\mathcal{V}\\ \|(f,g)\|_2=1}} \operatorname{Re}\left(\operatorname{Res}_0 \frac{ufg}{z^m}\right),$$

where $\mathcal{P}_{m-1} = H^2 \ominus z^m H^2$, the subspace of H^2 consisting of polynomials of degree $\leq m-1$.

Example 4.7 Consider the $(m+1) \times (m+1)$ lower triangular Toeplitz matrix

$$T = \begin{pmatrix} c_0 & 0 & \dots & 0 \\ c_1 & c_0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ c_m & c_{m-1} & \dots & c_0 \end{pmatrix},$$

which represents the compression of the Toeplitz operator with symbol

$$u(z) = c_0 + c_1 z + \dots + c_m z^m$$

to $H^2 \ominus z^{m+1} H^2$. In this case,

$$[Tf, g] = \frac{(ufg)^{(m)}(0)}{m!}$$

where the superscript (m) denotes the mth derivative of the given function. This yields the formula

$$\| T \| = \max_{\substack{\deg f \leq m \\ \| f \|_2 = 1}} \operatorname{Re} \frac{(uf^2)^{(m)}(0)}{m!} \tag{4.8}$$

for the norm of T. In particular, there exists an analytic function F on $\mathbb{D}$ such that $F(z) = c_0 + c_1 z + \dots + c_m z^m + O(z^{m+1})$ and $\| F \|_\infty \leq M$ if and only if (4.8) is $\leq M$ (see [1, 22]).

Example 4.8 Suppose that φ is a finite Blaschke product (4.5) with m distinct zeros $\lambda_1, \dots, \lambda_m$ and let u be a rational function which is bounded on $\partial\mathbb{D}$. Letting T denote the compression of the Toeplitz operator T_u to $H^2 \ominus \varphi H^2$, we find that (using Lemma 4.3)

$$\begin{aligned} [Tf, g] &= \frac{1}{2\pi} \int_{-\pi}^{\pi} u(e^{it}) f(e^{it}) \overline{Cg(e^{it})}\, dt \\ &= \frac{1}{2\pi} \int_{-\pi}^{\pi} u(e^{it}) f(e^{it}) g(e^{it}) \overline{\varphi(e^{it})} e^{it}\, dt \\ &= \frac{1}{2\pi i} \int_{\partial\mathbb{D}} \frac{ufg}{\varphi}\, d\zeta, \end{aligned}$$

which can be computed via the residue theorem. For instance, if u is analytic on $\mathbb{D}$, then (since the zeroes of φ are simple) this reduces to

$$[Tf, g] = \sum_{j=1}^{m} \frac{u(\lambda_j) f(\lambda_j) g(\lambda_j)}{\varphi'(\lambda_j)}$$

whence, for instance,

$$\| T \| = \max_{\substack{f \in H^2 \ominus \varphi H^2 \\ \| f \| = 1}} \operatorname{Re} \sum_{j=1}^{m} \frac{(uf^2)(\lambda_j)}{\varphi'(\lambda_j)}.$$

Using the preceding examples as a guide, one can easily produce the corresponding bilinear forms in a wide variety of circumstances. Indeed, the expression

$$[Tf, g] = \frac{1}{2\pi i} \int_{\partial\mathbb{D}} \frac{ufg}{\varphi}\, d\zeta$$

for the bilinear form generated by the compression of T_u to $H^2 \ominus \varphi H^2$ is valid for any inner function φ and symbol u in $L^\infty(\partial\mathbb{D})$. For instance, the common situation where $u = \psi_1/\psi_2$ factors as the quotient of two inner functions ψ_1 and ψ_2 is easily handled since

$$[Tf, g] = \frac{1}{2\pi i} \int_{\partial\mathbb{D}} \frac{\psi_1 fg}{\psi_2 \varphi}\, d\zeta$$

can be computed, in principle, via the residue theorem and suitable limiting arguments.

Acknowledgements The third named author was partially supported by the National Science Foundation Grant DMS-0350911.

References

[1] J. Agler and J. McCarthy, Pick Interpolation and Hilbert Function Spaces (American Mathematical Society, Providence, RI, 2002).
[2] R. Coifman and Y. Meyer, Au-delà des Opérateurs Pseudo-différentiels, Astérisque Tome 57 (Société Mathématique de France, Paris, 1978).
[3] P. Davis, Interpolation and Approximation (Dover, New York, 1975).
[4] J. Danciger, A min-max theorem for complex symmetric matrices, Linear Algebra Appl. **412**, 22–29 (2006).
[5] P. L. Duren, Theory of H^p Spaces (Academic Press, New York, 1970).
[6] K. Friedrichs, On certain inequalities and characteristic value problems for analytic functions and for functions of two variables, Trans. Amer. Math. Soc. **41**, No. 3, 321–364 (1937).
[7] F. R. Gantmacher, The Theory of Matrices, Vol. 2 (Chelsea, New York, 1989).
[8] S. R. Garcia, Conjugation and Clark operators, Contemp. Math. **393**, 67–112 (2006).
[9] S. R. Garcia and M. Putinar, Complex symmetric operators and applications, Trans. Amer. Math. Soc. **358**, 1285–1315 (2006).
[10] S. R. Garcia and M. Putinar, Complex symmetric operators and applications II, Trans. Amer. Math. Soc. **359**, 3913–3931 (2007).
[11] P. R. Halmos, A Hilbert Space Problem Book, 2nd edition (Springer-Verlag, New York, 1982).
[12] R. A. Horn and C. R. Johnson, Matrix Analysis (Cambridge Univ. Press, Cambridge, 1985).
[13] P. Koosis, Introduction to H_p Spaces (Cambridge University Press, Cambridge, 1998).
[14] S. Krantz, S. Y. Li, P. Lin, and R. Rochberg, The effect of boundary regularity on the singular numbers of Friedrichs operators on Bergman spaces, Michigan Math. J. **43**, No. 2, 337–348 (1996).
[15] P. Lin and R. Rochberg, On the Friedrichs operator, Proc. Amer. Math. Soc. **123**, No. 11, 3335–3342 (1995).
[16] N. K. Nikolski, Operators, Functions, and Systems: An Easy Reading, Vol. I, Mathematical Surveys and Monographs Vol. 92 (American Mathematical Society, Providence, RI, 2002).
[17] N. K. Nikolski, Operators, Functions, and Systems: An Easy Reading, Vol. II, Mathematical Surveys and Monographs Vol. 93 (American Mathematical Society, Providence, RI, 2002).
[18] J. Peetre and R. Rochberg, Higher order Hankel forms, Contemp. Math. **185**, 283–306 (1995).
[19] M. Putinar and H. S. Shapiro, The Friedrichs operator of a planar domain, in: Complex Analysis, Operators, and related Topics, Operator Theory: Advances and Applications Vol. 113 (Birkhäuser, Basel, 2000), pp. 303–330.
[20] M. Putinar and H. S. Shapiro, The Friedrichs operator of a planar domain II, in: Recent Advances in Operator Theory and related Topics. The Béla Szőkefalvi-Nagy Memorial Volume, Proceedings of the Memorial Conference, Szeged 1999, Operator Theory: Advances and Application Vol. 127 (Birkhäuser, Basel, 2001), pp. 519–551.
[21] H. S. Shapiro, Some inequalities for analytic functions integrable over a plane domain, in: Proceedings of the International Conference on Approximation and Function Spaces, Gdansk 1979 (North-Holland, 1981), pp. 645–666.
[22] T. Takagi, On an algebraic problem related to an analytic theorem of Carathéodory and Fejér and on an allied theorem of Landau, Japan J. Math. **1**, 83–93 (1925).