

A COMPUTATIONAL TRIGONOMETRY, AND RELATED CONTRIBUTIONS BY RUSSIANS KANTOROVICH, KREIN, KAPORIN

K. E. GUSTAFSON

University of Colorado, Boulder, CO, USA

e-mail: gustafs@@euclid.colorado.edu

Общая тригонометрия для матриц и операторов, созданная, в основном, автором данной статьи, оказывается полезной при анализе вычислительных методов и в математической физике. Цель данной статьи — краткий обзор этой малоизвестной теории и демонстрация последних достижений и будущих направлений развития. Особое внимание уделяется обсуждению вклада русских математиков Л. Канторовича, М. Крейна и Л. Капорина.

1. Abstract operator semigroups

Approximately 30 years ago (1967) this author was working in the abstract theory of operator semigroups (Hille—Yosida theory). Given A an infinitesimal generator of a contraction semigroup W_t on a Banach space X , and B a bounded multiplicative perturbation, the following result was obtained. Recall that an operator T is strongly accretive if there exists some $m > 0$ such that $\operatorname{Re} \langle Tx, x \rangle \geq m \|x\|^2$ for all x in domain of T . Recall also that the infinitesimal generators are always dissipative, $\operatorname{Re} \langle Ax, x \rangle \leq 0$, i. e., $-A$ is accretive. Here $\langle x, y \rangle$ denotes any semi-innerproduct on the Banach space X .

Theorem 1. *If A is a generator and B is bounded and strongly accretive, then BA is a generator iff BA is dissipative. In particular, for bounded strongly accretive A and B on a Hilbert space, BA is accretive when*

$$\sin \phi(B) \leq \cos \phi(A), \quad (1)$$

where

$$\sin \phi(B) = \inf_{-\infty < \varepsilon < \infty} \|\varepsilon B - I\|, \quad (2)$$

and

$$\cos \phi(A) = \inf_{x \neq 0} \frac{\operatorname{Re} \langle Ax, x \rangle}{\|Ax\| \|x\|}. \quad (3)$$

The new entities $\sin \phi(B)$ and $\cos \phi(A)$ ushered in a new operator trigonometry. Originally these were just called $\sin B$ and $\cos A$ but the above notation is better because it emphasizes the operator angle $\phi(A)$: the maximum angle through which an operator A turns vectors.

Recently in the two books [1, 2] I have summarized what has been discovered (up to 1996) for the operator trigonometry. However, to continue the development of this paper it is useful to collect a few key results of the early operator trigonometry which we will need in the following. Two important early results were the Minmax Theorem, and also the Euler Equation following below.

Theorem 2 (1968). *For any strongly accretive bounded operator A on a Hilbert space X , one has*

$$\sup_{\|x\|=1} \inf_{-\infty < \epsilon < \infty} \|(\epsilon A - I)x\|^2 = \inf_{-\infty < \epsilon < \infty} \sup_{\|x\|=1} \|(\epsilon A - I)x\|^2. \quad (4)$$

This Minmax Theorem establishes that (2) deserves to be called $\sin \phi(B)$, i. e., for strongly accretive B one has the (essential for a trigonometry relation) $\sin^2 \phi(B) + \cos^2 \phi(B) = 1$.

Theorem 3 (1969). *For any strongly accretive bounded operator A on a Hilbert space X , the antieigenvalue functional*

$$\mu = \frac{\operatorname{Re} \langle Ax, x \rangle}{\|Ax\| \|x\|} \quad (5)$$

has Euler equation

$$2\|Ax\|^2 \|x\|^2 (\operatorname{Re} A)x - \|x\|^2 \operatorname{Re} \langle Ax, x \rangle A^* Ax - \|Ax\|^2 \operatorname{Re} \langle Ax, x \rangle x = 0. \quad (6)$$

When A is selfadjoint or normal, the Euler equation is satisfied not only by the first antieigenvectors of A , but also by all eigenvectors of A .

The term antieigenvalue was introduced in [3]. The Euler equation for the antieigenvalue functional may be considered as a significant extension of the Rayleigh – Ritz variational theory of eigenvectors of selfadjoint operators A . Not only are maximal stretchings included, but also maximal turnings are now included. The maximal stretchings occur at the eigenvectors of selfadjoint A , at which $\mu = 1$. The maximal turning angle $\phi(A)$ occurs at the first antieigenvectors, for which $\mu = \cos \phi(A)$.

Specializing further to A a strongly positive selfadjoint operator, I determined in the early period 1966–1969 that

$$\cos \phi(A) = \frac{2\sqrt{mM}}{m+M}, \quad \sin \phi(A) = \frac{M-m}{M+m}, \quad (7)$$

where m and M are the lower and upper spectral bounds of A . Specializing further to A a symmetric positive definite (SPD) $n \times n$ matrix, the expressions (7) become

$$\cos \phi(A) = \frac{2\sqrt{\lambda_1 \lambda_n}}{\lambda_1 + \lambda_n}, \quad \sin \phi(A) = \frac{\lambda_n - \lambda_1}{\lambda_n + \lambda_1}. \quad (8)$$

These entities are achieved at the first antieigenvectors

$$x_{\pm}^1 = \pm \left(\frac{\lambda_n}{\lambda_1 + \lambda_n} \right)^{1/2} x_1 + \left(\frac{\lambda_1}{\lambda_1 + \lambda_n} \right)^{1/2} x_n. \quad (9)$$

Here λ_1 and λ_n denote the smallest and largest eigenvalues of A , x_1 and x_n corresponding eigenvectors, and all of x_1, x_n , and the pair x_{\pm}^1 have been normalized to norm 1.

The criteria (1), although not necessary, is surprisingly sharp. For example for $M_A = M_B = 1$ and $m_A = 1/2$, BA is accretive (i. e., $\operatorname{Re} BA$ is positive) whenever $m_B > 0.03$.

2. Krein's deviation

M. G. Krein in [20] introduced a quantity, which he called the deviation of A , $\text{dev}(A)$, which is equivalent to my operator angle $\phi(A)$. As I did, Krein uses the real angle: I had also defined an imaginary angle and a total angle, but these are less useful than the real angle. Through his function $am(A) = \min \text{dev}(\xi A)$, $|\xi| = 1$, Krein treated bounded operators A on a complex Hilbert space: I had treated arbitrary A in a Banach space, but the most important case is A bounded on a complex Hilbert space. Krein's work appeared slightly later than mine, but I assume it to be independent. Like me, Krein was motivated from a semigroups question, but it was a different question.

The principal motivation for Krein's introduction of the deviation $\text{dev}(A)$ was to bound the spectrum of an operator in a sector. This came about in treating an initial value problem

$$\begin{cases} \frac{dx(t)}{dt} = A(t)x(t) \\ x(0) = x_0 \end{cases} \quad (10)$$

where $A(t)$ is a time dependent infinitesimal generator found as the derivative of a function $F(t)$ of strongly bounded variation. Then the integral

$$V(t) = \int_0^t \exp\{dF(\tau)\} \quad (11)$$

solves the initial value problem with solution $x(t) = V(t)x_0$. Since $V(t)$ is in fact a multiplicative integral due to the multiplicative property of the exponentials, written formally

$$V = \Pi_0^t \exp dF \quad (12)$$

then one has formally

$$\text{dev}(V) \leq \sum_0^t \text{dev} \exp dF. \quad (13)$$

Since $\sigma(V)$ is in a sector bounded by $\text{dev}(V)$, this guarantees for example that the negative real axis is not in $\sigma(V)$, a useful property when integrating systems of ordinary differential equations.

To my knowledge, Krein did not develop further his $\text{dev}(A)$ theory. Nor did he introduce any antieigenvalue notion or theory. However, very nice convexity techniques were used by Mirman [22] to treat the notion of higher antieigenvalues $\mu_n(A)$ which I had also introduced in [3]. Beyond Krein and Mirman, I do not know of any further work on the antieigenvalue theory by the Russian school.

It should be mentioned that the German matrix specialist H. Wielandt also introduced a notion essentially equivalent to my operator angle $\phi(A)$, and at about the same time. He called his turning angle, the matrix singular angle. This idea appeared only in his lecture notes and not in the open literature and I was not aware of it until 1994 when an editor (H. Schneider) asked me to write a review [4] of it. More information about Wielandt's motivation (spectral inequalities of Weyl—Lidskii [21] type) may be found in the review [4].

3. Kantorovich's bound

Being invited to speak at the 1990 conference [5], I decided to take that opportunity to examine the possible connections of the operator trigonometry to computational linear algebra. I had suspected such connections 30 years ago, I even mentioned it in one of the 1968 papers. Very quickly the following result was obtained.

Theorem 4 (1990). *In quadratic steepest descent solution of the symmetric positive definite linear matrix system $Ax = b$, the fundamental Kantorovich error bound [17]*

$$E_A(x_{k+1}) \leq \left(1 - \frac{4\lambda_1\lambda_n}{(\lambda_1 + \lambda_n)^2}\right) E_A(x_k) \quad (14)$$

is in fact trigonometric:

$$E_A(x_{k+1}) \leq \sin^2 \phi(A) E_A(x_k). \quad (15)$$

Here E_A denotes the energy error inner product $E_A(x) = \langle (x-x^*), A(x-x^*) \rangle$, where x^* is the true solution of the system. Steepest descent is asymptotically a very slowly converging scheme but its analysis opens the way to that of better gradient schemes such as Conjugate Gradient, GMRES, others. In hindsight, I had independently derived my own version of the Kantorovich theory, 30 years ago, when I obtained the expressions (7) from convexity arguments on norms. In any case, the conclusion (15) from (14) was immediate, given the Minmax Theorem from the operator trigonometry. But until I put the two theories together, no one had seen the natural geometry (15) of the longstanding 1948 [17] Kantorovich bound (14). This result first appeared in the Proceedings [5], which were published in spite of the war breaking out in Yugoslavia only a few months after the June 1990 conference in Dubrovnik.

There followed three papers [6–8], respectively, occasioned by the First Conference on the Numerical Range and Numerical Radius, the special LAA volume for Chandler Davis, and the special LAA volume for John Maybee. My interest in connecting the operator trigonometry to computational linear algebra grew and when invited to speak at the 1996 conference [9], I decided I should also look at possible connections of the operator trigonometry to general iterative operator splitting schemes in computational linear algebra. Very basic to such schemes is the Richardson method $x_{k+1} = x_k + \alpha(b - Ax_k)$ with iteration matrix $G_\alpha = I - \alpha A$ and where one chooses the parameter α to produce an optimal convergence rate.

Theorem 5 (1996). *In Richardson iterative solution of $Ax = b$ for strictly accretive A , the optimal parameter α is*

$$\alpha = \epsilon_m = \frac{\operatorname{Re} \langle Ax_\pm^1, x_\pm^1 \rangle}{\|Ax_\pm^1\|^2} \quad (16)$$

where ϵ_m is the minimizing parameter for $\sin \phi(A)$ in (2) and where x_\pm^1 are A 's first antieigenvalues (9). The optimal convergence rate of the Richardson scheme is $\sin \phi(A)$.

The operator trigonometry-computational linear algebra connection has recently been extended to Preconditioned Conjugate Gradient, Jacobi, Gauss – Seidel, SOR, SSOR, Uzawa, AMLI, ADI, Multigrid, Domain Decomposition, and related iterative solution methods for $Ax = b$. More details may be found in [10–12] and work in progress. To date the computational trigonometry has provided a new geometrical convergence theory for each of these schemes. Within multilevel methods a fundamental connection between the trigonometric antieigenvalues of this paper and the strengthened C.B.S. constants of the domain decomposition methods has been established. Superlinear convergence of conjugate gradient methods is seen in terms of higher antieigenvalue behavior. It is hoped that in the future the computational trigonometric

viewpoint may bring about new algorithms and improved convergence theory for already good algorithms such as GMRES, BICGSTAB.

4. Kaporin's condition number

Preconditioned conjugate gradient methods for the general solution of the matrix equation $Ax = b$ are becoming more important, especially as we model larger physical problems in three space dimensions. These methods have the merits of ease of implementation and, for very large problems, lower memory requirements. For their use as linear solvers of the discretization of partial differential equations in three dimensions, they appear to have a distinct time advantage ($O(n^{1.3})$ versus $O(n^{2.3})$) over direct solvers, as well. A central question for such iterative PCG methods is to estimate the number of iterations k required to reduce the relative error to ϵ after k iterations. Previously, such estimates were usually given in terms of the standard condition number $\kappa(A) = \lambda_{\max}/\lambda_{\min}$ ratio of the largest and smallest eigenvalues of A .

I. E. Kaporin [18, 19] introduced another condition number $\beta(A)$ for this purpose, based on the ratio of the trace and determinant of the matrix A . This estimate has in principle a disadvantage of requiring knowledge of all eigenvalues but has on the other hand an advantage of demonstrating the superlinear convergence often experienced in PCG applications. Kaporin's condition number for an $n \times n$ matrix A with n positive eigenvalues is defined to be

$$\beta(A) = \left(\frac{\sum_{i=1}^n \lambda_i}{n} \right)^n / \left(\prod_{i=1}^n \lambda_i \right). \quad (17)$$

The connection between β of (17) and μ from (5) and (8) is immediate for $n = 2$: $\beta^{-1/2} = \mu$. Thus it is useful to view the antieigenvalue μ as a condition number and to compare its potential uses to that of the conventional condition number κ and Kaporin's condition number β . One may easily verify their general relation

$$1 \leq \beta(A)^{1/n} \leq \kappa(A) \leq [\kappa(A)^{1/2} + \kappa(A)^{-1/2}]^2 = 4\mu(A)^{-2} \leq 4\beta(A). \quad (18)$$

As I pointed out in [9], the preconditioning strategies according to κ , β , and μ , are equivalent in practice but different in philosophy. In seeking

$$\kappa(BA) \sim 1 \quad (19)$$

we are seeking to "undilate" A with B . In seeking

$$\mu(BA) \sim 1 \quad (20)$$

we are seeking to "untwist" A with B . In seeking

$$\beta(BA) \sim 1 \quad (21)$$

we are trying to both undilate and untwist A at the same time. Perhaps it is better to say that with β we are trying to "unmoment" A with B .

In principle the use of $\mu(A)$ rather than $\beta(A)$ could have two advantages. First, in a number of matrix iterative computations, one may pull out estimates of the largest and smallest eigenvalues much easier than estimates of all eigenvalues. Second, $\mu(A)$ provides a clear geometrical picture, whereas $\beta(A)$ was developed out of log norm estimates and BFGS update strategies of quasi-Newton iterative schemes. In any case, $\mu(A)$ now brings to $\beta(A)$ a previously absent geometrical theory.

5. An Extended operator trigonometry

Originally the theory discussed here was stated in a very general context, A an unbounded operator in a Banach space, and also the immediately apparent vagaries about real, imaginary, and complex operator angles were taken care of simply by defining real, imaginary, and total cosines. Because the original semigroup theory motivation dealt with accretive and dissipative operators, therefore I emphasized the real operator trigonometry. Also, one may easily rotate any desired half-plane theory to the real theory.

However, in computational linear algebra, important applications and hence interest has recently been turning to the case of A a general, nonsymmetric, perhaps sparse, perhaps very large, matrix, often $n \times n$, invertible, and perhaps with only real entries. Having these applications in the back of my mind, for the last couple of years I have thought about the most natural way to extend the operator trigonometry to arbitrary matrices, and my opinion now is that one should use polar form. There are two strong contributing reasons for this. First, I have never been motivated in the operator trigonometry to think of uniformly turning operators A , e.g., those which rotate all vectors by a fixed angle. What interests us in the operator trigonometry is the relative turning of vectors, just as in the classical eigenvalue theory we are interested in the relative stretching of vectors. Thus polar form $A = U|A|$ efficiently removes the “uniform turning,” e.g., in U , and we already have the operator angle theory for $|A|$. Second, for invertible operators A , polar form is better than singular value decomposition for our purposes of an extended operator trigonometry, because we can show that the essential Minmax Theorem extends without incident.

Therefore let us change the key definition (2) to

$$\sin \phi(A) = \inf_{\epsilon} \|\epsilon A - U\|. \quad (22)$$

Of course in the A symmetric positive definite case, U is just the Identity. Then, considering first for example A to be an arbitrary $n \times n$ nonsingular matrix with singular values $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_n > 0$, we obtain from (22) and the second expression in (8) that

$$\sin \phi(A) = \min_{\epsilon > 0} \|\epsilon |A| - I\| = \frac{\sigma_1(A) - \sigma_n(A)}{\sigma_1(A) + \sigma_n(A)}. \quad (23)$$

One may check that the key min-max identity (4) in its essential form $\sin^2 \phi(A) + \cos^2 \phi(A) = 1$ is then satisfied if one modifies definition (3) to

$$\cos \phi(A) = \inf_{x \neq 0} \frac{\langle |A|x, x \rangle}{\||A|x\|| \|x\|}. \quad (24)$$

Then $\cos \phi(A)$ is given as in (8) with λ_1 and λ_n replaced by σ_n and σ_1 , respectively. Let us therefore formalize this result.

Theorem 6 (1998). *Let A be an arbitrary invertible operator on a complex Hilbert space X . From $A = U|A|$ polar form define the angle $\phi(A)$ and $\sin \phi(A)$ according to (22), and $\cos \phi(A)$ according to (24). Then*

$$\sin^2 \phi(A) + \cos^2 \phi(A) = 1 \quad (25)$$

and a full operator trigonometry of relative turning angles obtains for A from that of $|A|$.

Proof. When A is invertible, the partial isometry U in A 's polar form is unitary and $|A|$ is strongly positive selfadjoint. Looking now at the right hand side of (4), we may write

$$\|(\epsilon A - U)x\|^2 = \|U(\epsilon|A| - I)x\|^2 = \|(\epsilon|A| - I)x\|^2 \tag{26}$$

from which

$$\min_{-\infty < \epsilon < \infty} \max_{\|x\|=1} \|(\epsilon A - U)x\|^2 = \min_{\epsilon > 0} \|\epsilon|A| - I\|^2. \tag{27}$$

Thus, in view of definition (22) and the known second expression in (7) for strongly positive selfadjoint operators, we obtain from (27) that $\sin^2 \phi(A) = \sin^2 \phi(|A|)$, and hence

$$\sin \phi(A) = \sin \phi(|A|) = \frac{\|A\| - \|A^{-1}\|^{-1}}{\|A\| + \|A^{-1}\|^{-1}}. \tag{28}$$

Recall that $\|A\| = M = \||A|\|$ and that $\|A^{-1}\|^{-1} = m = \||A|^{-1}\|^{-1}$ and note that we prefer the norm expressions in (28) rather than the m and M from $|A|$ because we speak of A in the Theorem, even though we may need to use $|A|$ to evaluate the expressions in (28). Continuing, by the definition of $\cos \phi(A) = \cos \phi(|A|)$ in (24), the left hand side of (4) becomes equal to the right hand side of (4), namely

$$1 - \cos^2 \phi(A) = 1 - \cos^2 \phi(|A|) = \sin^2 \phi(A). \tag{29}$$

Here we have made use of the Minmax Theorem applied to the strongly accretive operator $|A|$, and we have also used the argument of [8, (3.10)] to get to the expression $1 - \cos^2 \phi(A)$. Thus we have shown (25). All spectral details known previously for the strongly positive selfadjoint operator trigonometry now transfer via $|A|$ to A .

Let us next consider the second principal theoretical result of the early operator trigonometry, namely the Euler equation (6) for the antieigenvalue functional (5). That is, let us insert $A = U|A|$ for A a strongly accretive (hence, invertible) bounded operator, into the Euler equation (6), from which we arrive at the expression in terms of the U and $|A|$ of A 's polar form,

$$(U|A| + |A|U^*)x - \frac{\langle (U|A| + |A|U^*)x, x \rangle}{2} \cdot \frac{|A|^2 x}{\||A|^2 x\|} - \frac{\langle (U|A| + |A|U^*)x, x \rangle}{2\|x\|^2} x = 0. \tag{30}$$

Apparently our approach to a general operator trigonometry based upon polar form has not produced an appetizing Euler equation (30).

However, when we remember that the Euler equation (6) was originally derived from the (real) antieigenfunctional (5) due to our earlier interest in accretive semigroup generators, whereas now the antieigenfunctional of interest is (24) as motivated now by the needed identity (25), we see that the appropriate Euler equation for the extended operator trigonometry is in fact before us, as follows.

Theorem 7 (1999). *For $A = U|A|$ an arbitrary invertible operator on a complex Hilbert space X , and normalizing solutions x to $\|x\| = 1$ for convenience, the appropriate Euler equation in this extended operator trigonometry is*

$$\frac{|A|^2 x}{\langle |A|^2 x, x \rangle} - \frac{2|A|x}{\langle |A|x, x \rangle} + x = 0. \tag{31}$$

Proof. Starting from the functional (24), because $|A|$ is strongly accretive, the derivation of the Euler equation of that functional is the same as the original derivation of (6) from (5). See

e. g. [1, 2] or [7]. Then the simplification to (31) in the selfadjoint case follows, e. g. as already observed in [7].

The Euler equation in the extended operator trigonometry of this paper is actually simpler than the original one, because we have given up the emphasis on general accretive operators, in going to polar form.

6. Future directions and remarks

6.1. Iterative methods

It is expected that the extended operator trigonometry of Section 5 will open large portions of the analysis of iterative computational methods for nonsymmetric matrices A to the same new trigonometric geometrical and convergence rate investigations available to date essentially only for symmetric matrices A , e. g. as in [9–12]. Further investigation of such computational trigonometry will be pursued elsewhere.

6.2. Alternate extensions

We are not asserting that our approach in this paper is the only way in which one may obtain an extended operator trigonometry. Indeed, see on earlier discussion [8, Remarks 4.1 and 4.2] where we point out some discrepancies between the variational formulation and the Euler formulation, such discrepancies having their origin in our original interest in and emphasis on real accretivity of operators. Also, there could be situations in which the singular value decomposition could yield an advantageous theory. Even within our approach, there may in the future be a need in certain applications to also take into account certain phase angles ψ within U , e. g., two-dimensional internal plane rotations as they occur in the Jacobi scheme for eigenvalue calculations, or situations in which the action of U is that of permutation or change of basis.

However, the theory we have presented here is very general. It applies to operators A as well as matrices A . For computational linear algebra purposes all spectral information for the operator trigonometry of an $n \times n$ matrix A may be obtained from its singular values σ_1 and σ_n . This dependence on estimates of singular values may be a cause of some computational complaint because it requires in principle approximately solving the eigenvalue problem for A^*A and the latter also requires a (possibly large scale) multiplication by A^* . On the other hand most practitioners in the computational linear algebra community when presented with a large nonsymmetric matrix A immediately think subconsciously about what its condition number $\kappa = \sigma_1/\sigma_n$ may be, because in principle that ratio is one's best initial feeling as to the sensitivity of $Ax = b$ to errors and machine roundoff. Thus some consideration and thinking about A 's conditioning, e. g., its singular value magnitudes, is unavoidable.

6.3. Condition number angle

An overpreoccupation with condition number $\kappa = \sigma_1/\sigma_n$ by the computational community is perhaps the reason that the true geometrical meaning of the Kantorovich bound [17] was not seen between 1948 and 1990 [5]. There was indeed an angle defined, the so-called Kantorovich–Wielandt condition number angle θ , defined by $\cot(\theta(A)/2) = \kappa$. I showed in [1, 2] that one may connect my antieigenvalue operator angle $\phi(A)$ to the condition number angle $\theta(A)$, in the

case of A symmetric positive definite, as follows: $\cos \phi(A^2) = \sin \theta(A)$. Here let us go further, by use of the extended operator trigonometry of Section 5, to state the following extension of that result to all nonsingular matrices.

Theorem 8 (1998). *For any nonsingular $n \times n$ matrix A , the condition number angle $\theta(A)$ and the operator angle $\phi(A)$ are related by*

$$\cos \phi(A^*A) = \sin \theta(A). \quad (32)$$

*The condition number angle $\theta(A)$ is determined by the first antieigenvectors of A^*A . Its $\sin \theta(A)$ is the first antieigenvalue of A^*A .*

6.4. Invertibility

We have taken A invertible here to mean A^{-1} also is in $\mathcal{B}(X)$ but it should be mentioned that much of the extended theory above holds as well for those instances in which A is just 1–1. There are three cases here (e. g. see [13]) for Hilbert space operators, corresponding to whether 0 is in the continuous or residual spectrum of A : A is 1–1 with dense range, range of A is not dense but A^{-1} is bounded, or range of A is not dense and A is only 1–1. The point is that U is an isometry when A is 1–1, and then it may be checked that (26) is still true. However, $|A|^{-1}$ may be bounded or unbounded.

The situation of A bounded but A not 1–1, e. g., a semidefinite matrix, is also of interest in the computational linear algebra applications. That extension of the theory of this paper will be pursued elsewhere. In that case U is a partial isometry and the null space of A agrees with that of $|A|$.

When A is an unbounded operator in a Hilbert space, an early result in the operator trigonometry (see [1, 2]) showed that $\cos \phi(A) = 0$. Because $\phi(A) = \phi(A^{-1})$ when A has an inverse, to date those instances in which A or A^{-1} is unbounded have not been interesting for computational algorithms.

6.5. Equivalent reformulation

Because $|A| = U^*A$, we may initially define the operator angle $\phi(A)$ from

$$\cos \phi(A) = \inf_{x \neq 0} \frac{\langle Ax, Ux \rangle}{\|Ax\| \|x\|} \quad (33)$$

and arrive at the same full extended trigonometry of Section 5.

6.6. Higher antieigenvectors

Originally, 30 years ago I defined higher antieigenvalues in terms of infima of the functional (5) restricted to subspaces orthogonal to the preceding antieigenvectors. In the 1990's my computer experiments with the iterative linear solvers for $Ax = b$ indicated that it was better to define higher antieigenvalues combinatorially, e. g., replacing λ_n and λ_1 in (9) by λ_j and λ_i , $1 < i, j < n$, and then defining the higher antieigenvectors by replacing the corresponding eigenvector components accordingly. One thereby “nests inward” through a sequence of decreasing critical turning angles. See for example the discussions in [7, Remark 6.2] and [8, Section 5]. Let us observe here that these two philosophies are not necessarily so different. Consider

any two vectors $x_+ = cx_1 + dx_2$ and $x_- = -cx_1 + dx_2$ where c and d are nonzero possibly complex coefficients and where x_1 and x_2 are linearly independent. Then x_+ and x_- have the same span as x_1 and x_2 . Thus minimizing the functional (5) on the subspace orthogonal to the span of the first antieigenvector pair is the same as finding the next critical turning angle of A in the subspace orthogonal to the span of the first and last eigenvector of A , at least in the cases (e. g., A symmetric or normal) where A 's antieigenvectors may be expressed in terms of A 's extreme eigenvectors as in (9).

Pursuing this analysis a bit further, now also assume x_1 and $x_2 \equiv x_n$ to have norm one. Then (using (9) and (8) respectively in the second and third lines therein) one obtains the interesting expression

$$\begin{aligned} \langle x_+, x_- \rangle &= -|c|^2 + |d|^2 + 2i\text{Im}(c\bar{d}\langle x_1, x_n \rangle) = \\ &= \frac{\lambda_1 - \lambda_n}{\lambda_1 + \lambda_n} + i \frac{2\lambda_1^{1/2}\lambda_n^{1/2}}{\lambda_1 + \lambda_n} \text{Im} \langle x_1, x_n \rangle = \\ &= -\sin \phi(A) + i \cos \phi(A) \text{Im} \langle x_1, x_n \rangle. \end{aligned} \quad (34)$$

In particular for A symmetric positive definite we have from (34) that $\langle x_+^1, x_-^1 \rangle = -\sin \phi(A)$, i. e., the angle ϕ_{\pm} of the antieigenvector pair (9) is always $\phi(A) + \pi/2$. This fact has not been heretofore mentioned. The angles ϕ_{\pm}^k of higher antieigenvector pairs defined combinatorially as described above will enjoy a similar relation in terms of the higher operator angles ϕ^k .

Although the antieigenvectors x_{\pm}^1 in (9) occur in pairs, their linear span does not form an anti-eigenspace. However, if the λ_1 or λ_n has multiplicity greater than one, then the respective x_1 or x_n in (9) may be taken arbitrarily of norm one from the respective eigenspace. This fact is easily verified but has not been pointed out heretofore. In this sense the first antieigenvectors take on the multiplicities of the first and highest eigenspaces. The same statement is valid for the higher antieigenvectors.

6.7. Other applications

The computational trigonometry presented in this paper also has interesting recent applications to control theory [14], wavelets [15], and quantum probability [16].

References

- [1] GUSTAFSON K. *Lectures on Computational Fluid Dynamics, Mathematical Physics, and Linear Algebra*. Kaigai Publications, Tokyo, 1996; World-Scientific, Singapore, 1997.
- [2] GUSTAFSON K., RAO D. *Numerical Range: The Field of Values of Linear Operators and Matrices*. Springer Verlag, New York, 1997.
- [3] GUSTAFSON K. Antieigenvalue inequalities in operator theory. In: *"Inequalities III"*, Los Angeles, 1969. Ed. O. Shisha, Academic Press, 1972, 115–119.
- [4] GUSTAFSON K. Commentary on Topics in the Analytic Theory of Matrices. In: *"Collected Works of Helmut Wielandt"*, 2. Eds. B. Huppert, H. Schneider. De Gruyters, Berlin, 1996, 356–367.
- [5] GUSTAFSON K. Antieigenvalues in analysis. Eds. C. Stanojevic, O. Hadzic. *Proc. Fourth Intern. Workshop in Analysis and its Applications*, Dubrovnik, 1990. Novi Sad, Yugoslavia, 1991, 57–69.

- [6] GUSTAFSON K. Operator trigonometry. *Linear and Multilinear Algebra*, **37**, 1994, 139–159.
- [7] GUSTAFSON K. Antieigenvalues. *Linear Algebra Appl.*, 208/209, 1994, 437–454.
- [8] GUSTAFSON K. Matrix trigonometry. *Ibid.*, 217, 1995, 117–140.
- [9] GUSTAFSON K. Trigonometric interpretation of iterative methods. Eds. O. Axelsson, B. Polman. *Proc. Conf. Algebraic Multilevel Iteration Methods with Appl.*, Nijmegen, Netherlands, 1996, 23–29.
- [10] GUSTAFSON K. Operator trigonometry of iterative methods. *Num. Lin. Alg. with Appls.*, **4**, 1997, 333–347.
- [11] GUSTAFSON K. Domain decomposition, operator trigonometry, Robin condition. *Contemporary Mathematics*, 218, 1998, 455–460.
- [12] GUSTAFSON K. Operator trigonometry of the model problem. *Num. Lin. Alg. with Appls.*, **5**, 1999 (to appear).
- [13] GUSTAFSON K. Operator spectral states. *Computers Math. Appls.*, **34**, 1997, 467–508.
- [14] GUSTAFSON K. Operator trigonometry of linear systems. Eds. N. Koussoulas, P. Groumpos. *Proc. IFAC Symp. on Large Linear Systems*, Patras, Greece, 1998, 950–955.
- [15] GUSTAFSON K. Operator trigonometry of wavelet frames. Eds. J. Wang, M. Allen, B. Chen, T. Mathew. *Proc. IMACS Conf. on Iterative Methods*, Jackson, Wyoming, 1997, IMACS Publications. New Brunswick, 1998, 161–166.
- [16] GUSTAFSON K. The geometry of quantum probabilities. In: “*On Quanta, Mind, and Matter. Hans Primas in Context*”. Eds. A. Amann, H. Atmanspacher, U. Mueller — Herold. Kluwer, Dordrecht, 1999, 151–164.
- [17] KANTOROVICH L. Functional analysis and applied mathematics. *Uspehi Mat. Nauk*, **3**, No. 6, 1948, 89–185.
- [18] KAPORIN I. E. An alternative approach to estimation of the conjugate gradient iteration number. In: “*Numerical Methods and Software*”. Eds. Yu. A. Kuznetsov, Acad. Sci. USSR, Department of Computational Mathematics, 1990, 53–72 (in Russian).
- [19] KAPORIN I. E. New Convergence results and preconditioning strategies for the conjugate gradient method. *Num. Lin. Alg. with Appls.*, **1**, 1994, 179–210.
- [20] KREIN M. G. Angular localization of the spectrum of a multiplicative integral in a Hilbert space. *Func. Anal. Appl.*, **3**, 1969, 89–90.
- [21] LIDSKII V. B. The proper values of the sum and product of symmetric matrices. *Doklady Akad. Nauk SSSR*, **75**, 1950, 769–772.
- [22] MIRMAN B. Antieigenvalues: method of estimation and calculation. *Linear Algebra Appl.*, **49**, 1983, 247–255.