Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 30, pp. 649-669, October 2015

ELA
http://repository.uwyo.edu/ela

A STRANGE PHENOMENON FOR THE SINGULAR VALUES OF COMMUTATORS WITH RANK ONE MATRICES*

DAVID WENZEL†

Dedicated to Dieter Happel.

This paper could have been presented on the memorable occasion of a great mathematician's 60th birthday. Sadly... we lost him the year before. Though I'm not personally active in his field of research and this publication is not related to his achievements, I want to express that we at the Chemnitz department of mathematics will miss our dean, colleague, and friend.

Abstract. The singular values of $XY - YX$ are the objects under investigation. Here, X and Y are square matrices with complex entries, and one of them has rank one. Hence, there are at most two non-trivial numbers among the commutator's singular values, and the pairs of interest can be depicted in the plane. The emphasis will lie on the unexpectedly intriguing case in which both matrices are of rank one – because the result then is astonishingly complex, and the problem gives rise to interpretations unveiling geometry acting in the background.

Key words. Commutator, Singular Values.

AMS subject classifications. 15A18.

1. Introductory words. In [2], questions concerning the magnitude of the commutator of two matrices were raised. Suppose $X, Y \in \mathbb{R}^{n\times n}$ with $\|X\|_{\mathsf{F}} = \|Y\|_{\mathsf{F}} = 1$, where $\|\cdot\|_{\mathsf{F}}$ denotes the Frobenius norm. It was shown that $\|XY - YX\|_{\mathsf{F}}$ tends to zero in the mean with increasing matrix size n under certain reasonable conditions on the underlying distribution. As an exemplification, if X and Y are drawn independently from the (n^2-1)-dimensional unit sphere, then the expected value and the variance of the random variable $\|XY - YX\|_{\mathsf{F}}^2$ are given by (see [2, Theorem 1.1])

$$E\big(\|XY - YX\|_{\mathsf{F}}^2\big) = \frac{2}{n} - \frac{2}{n^3} \leq \frac{2}{n} \quad \text{and} \quad V\big(\|XY - YX\|_{\mathsf{F}}^2\big) \approx \frac{8}{n^4}. \tag{1.1}$$

*Received by the editors on June 17, 2013 Accepted for publication on February 16, 2015, Handling Editor: Bryan Shader.

†Department of Mathematics, Chemnitz University of Technology, Chemnitz, Germany (david.wenzel.mathematik@t-online.de).

To this effect, the norm of the commutator of two large matrices is typically quite small – though almost surely nonzero.

The conjecture

$$\|XY - YX\|_{\mathsf{F}} \leq \sqrt{2}\|X\|_{\mathsf{F}}\|Y\|_{\mathsf{F}} \tag{1.2}$$

was posed (see [2, Conjecture 1.2]) and was proven independently and differently in [6] and [7]. Pretty much the same time, (1.2) was shown to be valid even for all complex $n \times n$ matrices ([3, Theorem 2.2]). The constant $\sqrt{2}$ in the upper bound is not trivial – the triangle inequality and the submultiplicativity only grant

$$\|XY - YX\|_{\mathsf{F}} \leq \|XY\|_{\mathsf{F}} + \|YX\|_{\mathsf{F}} \leq 2\|X\|_{\mathsf{F}}\|Y\|_{\mathsf{F}}. \tag{1.3}$$

The worst-case-relation was studied with other norms (esp. these of the Schatten type). The analogous inequality was then determined in many cases via a modified Riesz-Thorin technique of complex interpolation; introduced in [8], this resulted in [9]. Moreover, it was proven that equality in (1.2) may be attained only for very special pairs of matrices ([3, Corollary 4.2]):

> *If X and Y form a maximal pair, i.e. X and Y are not zero and $\|XY - YX\|_{\mathsf{F}} = \sqrt{2}\|X\|_{\mathsf{F}}\|Y\|_{\mathsf{F}}$, then*
>
> $$\operatorname{rank} X \leq 2, \quad \operatorname{rank} Y \leq 2, \quad \operatorname{trace} X = 0, \quad \operatorname{trace} Y = 0,$$
>
> *and, in the Frobenius inner product, X and Y are orthogonal to each other, as well as to several matrices built from these.*

We now return to the problem of the average case. Although the maximal value is known to be rare, this is not a sufficient explanation of the pretty small mean value seen in (1.1). Fig. 1.1 gives an illustration and compresses the initial observations of [2, Fig. 1] into one picture. The aim is to get some more insight into the structure of the commutator operation in order to find a reason for the behaviour... and to provide an answer to the question as to why $XY - YX$ cannot be expected to have large norm. Our investigation of the dependence of the singular values of $XY - YX$ on the matrices X and Y will be more detailed than the usual norm bounds. This results into specific sets that turn out to be smaller than one could have thought. Here, we only treat the situation in which one of the matrices X and Y has rank one. In most cases we will obtain an intuitive triangle (Theorem 3.2). Whereas this has its origins in suitable norm estimates, pairs of two rank one matrices behave differently and surprise with a strange shape (Conjecture 3.6, Theorem 4.1). There is a much more complicated, geometric reason that is not easy to handle. As this article is part of the *Electronic Journal of Linear Algebra*, we are able to profit of multimedia features. Selected figures are made interactive, giving a better understanding. Depending on your pdf viewer, some features may be unavailable. It is best using Acrobat.

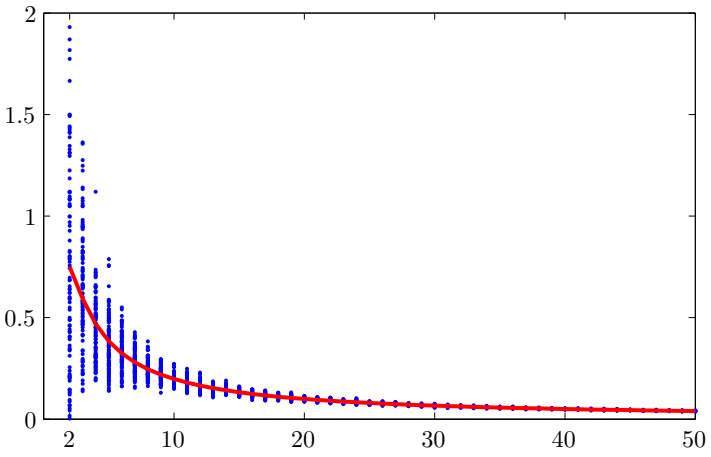


FIG. 1.1. *Norm of the commutator in different dimensions.*
$\|XY - YX\|_{\mathsf{F}}^2$ is plotted over the size n of X and Y, with values for 100 random matrix pairs each (normalised). The predicted trend in the form of the expected value (1.1) is charted as a red curve. The small variance is reflected in the points' clustering within a band getting more and more tight with growing size. Its width is of order $\frac{1}{n^2}$, the standard deviation.

Note: Depending on your pdf viewer, some features may be unavailable. It is best using Acrobat.

By O_n we denote the $n \times n$ zero matrix, and I_n stands for the identity of size $n \times n$. We use the notation $M \oplus N$ for the block diagonal matrix constructed of M and N. For any matrix M, the adjoint is M^*, and the vector of (decreasingly ordered) singular values $s_i = s_i(M)$ will be denoted by $\sigma(M) = (s_1, \ldots, s_n)$. The already met Frobenius norm can be represented in two way, either with help of the entries m_{jk} or in terms of s_i: $\|M\|_{\mathsf{F}}^2 = \sum |m_{jk}|^2 = \sum s_i^2$. It is part of the family of Schatten p norms $\|M\|_p^p = \sum s_i^p$ $(1 \le p \le \infty)$. In this sense, $\|\cdot\|_{\mathsf{F}}$ is $\|\cdot\|_2$, and $\|\cdot\|_\infty$ gives the widely known spectral norm. Here, clarification is wise since the last one often gets subscript 2, too. Later on, $\|\cdot\|$ is written for the usual Euclidean vector norm.

2. How to capture the shape of a bilinear matrix map. Regard the commutator as a function $(X, Y) \mapsto XY - YX$ that takes two matrices and turns them into another one. Our attempt to catalogue some properties of such an abstract function is to determine the range of the image's singular value vector where those of both arguments are given in accordance with a special scheme.

DEFINITION 2.1. A $n \times n$ matrix $M \in \mathbb{C}^{n\times n}$ will be said to be of *strict rank* r if

$$\sigma(M) = (\underbrace{s, \ldots, s}_{r}, \underbrace{0, \ldots, 0}_{n-r})$$

with $s > 0$; we abbreviate this notion by $M \sim^{(r)} s$.

Let $a, b \in \mathbb{C}$. For $f : \mathbb{C}^{n\times n} \to \mathbb{C}^{n\times n}, (X, Y) \mapsto aXY + bYX$, the sets

$$\mathcal{S}_{\tilde{r},r}(f) = \left\{\sigma(f) \ : \ X \sim^{(\tilde{r})} \tfrac{1}{\sqrt{\tilde{r}}},\ Y \sim^{(r)} \tfrac{1}{\sqrt{r}}\right\}$$

together form its σ-atlas.

Obviously, all rank one matrices are strict, but this is not always true for $r \geq 2$. If $Y \sim^{(r)} \frac{1}{\sqrt{r}}$, then $\|Y\|_{\mathsf{F}} = 1$. So, a useful norm assumption based on the homogenicity of singular values is already incorporated into the convention above.

When investigating $XY - YX$, $\tilde{r} \leq r$ may be demanded due to symmetry. A bit sloppy, we also suppress the matrix order in what follows. At this point, we are hinting that (in the case of the commutator) matrices are expected to be small, and large ones presumably don't contribute new information to the problem's solution. Every matrix can be padded with zeros to match larger sizes. Anyhow, X and Y have to be at least $r \times r$.

This publication considers only the cases with $\tilde{r} = 1$. As $\operatorname{rank} X = 1$ implies $\operatorname{rank} XY = \operatorname{rank} YX = 1$, the commutator then has rank less or equal to two. Its singular values (excluding the last $n - 2$ trivial ones) can hence be drawn in a plane. Because of the non-negativity of singular values and the chosen order $s_1 \geq s_2 \geq \cdots$, we may also restrict ourselves to a sector with angle $\frac{\pi}{4}$. The simplified notation

$$\mathcal{S}_r = \{(s_1, s_2) \ : \ (s_1, s_2, 0, \ldots, 0) \in \mathcal{S}_{1,r}(XY - YX)\}$$

is introduced to make it easier to study parts of the σ-atlas with the help of planar graphs. Assigning the properties to X and Y in that order is no loss of generality.

Known results on singular values of arbitrary matrices rarely give exact information, but they can be used to limit the regions we look for. The singular spectrum of a 2×2 matrix M can be calculated explicitly via

$$s_{1,2}^2(M) = \frac{\|M\|_{\mathsf{F}}^2}{2} \pm \sqrt{\frac{\|M\|_{\mathsf{F}}^4}{4} - |\det M|^2}. \tag{2.1}$$

For general $M, N \in \mathbb{C}^{n\times n}$, Ky Fan established estimates for the singular values of products and sums; also Weyl is credited for the result. Writing $\sigma(M) = (\mu_1, ..., \mu_n)$ and $\sigma(N) = (\nu_1, ..., \nu_n)$, M and N suffice the following inequalities ([1],[5]):

$$\max_{i+j=k+n} \mu_i\nu_j \leq \quad s_k(MN) \quad \leq \min_{i+j=k+1} \mu_i\nu_j, \tag{2.2}$$

$$\max_{i-j=k-1} \{0, \mu_i - \nu_j, \nu_i - \mu_j\} \leq \ s_k(M + N) \ \leq \min_{i+j=k+1} \mu_i + \nu_j. \tag{2.3}$$

That the estimates are not tight is demonstrated by the next example.

EXAMPLE 2.2. Take a closer look at a rank one matrix X. Y is arbitrary. Let $\|X\|_{\mathsf{F}} = \|Y\|_{\mathsf{F}} = 1$. As pointed out before, the choice of X guarantees that the commutator's singular values $s_1 \geq s_2 \geq 0$ can be depicted.

With the usual norm bounds, we are able to restrict a combination of all singular values. For instance, (1.3) implies that the pairs of the two largest ones are located inside a circle of radius 2. The same procedure applies to the Schatten ∞ norm:

$$\|XY - YX\|_\infty \leq 2\|X\|_\infty \|Y\|_\infty. \tag{2.4}$$

Hence, the largest singular value $s_1(XY - YX) = \|XY - YX\|_\infty$ is bounded by 2. This is a weaker statement, but it inspires to think about estimating s_2 and turns our attention to the Ky Fan inequalities.

By (2.2), we obtain $0 \leq s_1(XY) \leq s_1(X)s_1(Y)$ (or by using $s_1(M) = \|M\|_\infty$ and certain norm inequalities: $0 \leq \|XY\|_\infty \leq \|X\|_\infty \|Y\|_\infty$). The other estimates yield $s_2(XY) = \cdots = s_n(XY) = 0$, whence the product retains rank one. Obviously, similar assertions can be deduced for YX. In addition, (2.3) gives

$$\begin{aligned} 0 \leq \big|s_1(XY) - s_1(YX)\big| &\leq s_1(XY - YX) \\ &\leq s_1(XY) + s_1(YX) \leq 2s_1(X)s_1(Y), \end{aligned} \tag{2.5}$$

$$\begin{aligned} 0 &\leq s_2(XY - YX) \\ &\leq \min\{s_1(XY), s_1(YX)\} \leq s_1(X)s_1(Y), \end{aligned} \tag{2.6}$$

and $s_3(XY - YX) \leq s_2(XY) + s_2(-YX) = 0$, ensuring the rank two property.

Clearly, (2.5) proves (2.4). The relation (2.6) for the second singular value is a new restriction, partially better than (1.3). Nevertheless, we meanwhile know about (1.2). This guarantees that the pair (s_1, s_2) doesn't lie outside the circle with radius $\sqrt{2}$, and gives the best bound, so far. A realisation of non-zero equality in Ky Fan's inequalities is barely possible with the very special case of the commutator. All bounds to the attainable singular value pairs can be compared to each other in Fig. 2.1.

REMARK 2.3. The estimates given in the last example depend solely on the rank one restriction to X, but the sharpest inequality is free of this requirement. We know that singular values of *any* commutator of normalised matrices are inside the $\sqrt{2}$-ball (a disc in the previous situation). Of course, as one may have expected, all these points do represent such a pair if X and Y are allowed to be arbitrary. Plotting the singular values of enough random 2×2 matrix pairs will cover the whole sector. Similar pictures for $n \times n$ matrices cannot be given that easy as there are up to n components in $\sigma(XY - YX)$. But, you may check that larger matrices result in smaller norms that actually bound the singular values (Fig. 1.1). The message: points are easier to find close to the origin and hardly reach the $\sqrt{2}$-sphere anymore.

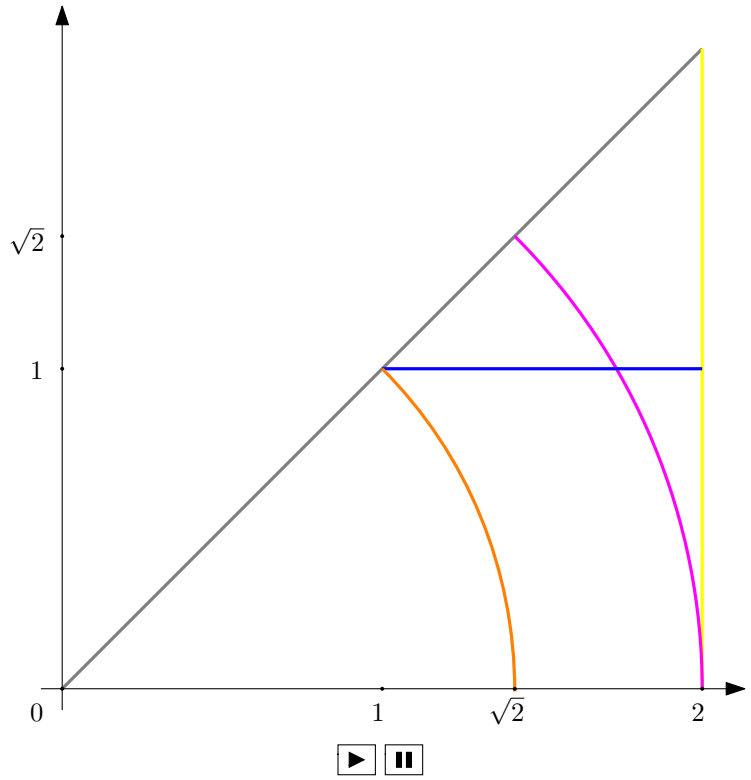


FIG. 2.1. *Bounds to the singular values of the commutator of norm 1 matrices.* Different inequalities restricting (s_1, s_2): ordering $s_1 \geq s_2$ (below grey line), trivial ∞-estimate (2.4) or Ky Fan-s_1-inequality (2.5) (left of yellow line), trivial F-estimate (1.3) (left of magenta arc), Ky Fan-s_2-inequality (2.6) (below blue line), sharp F-estimate (1.2) (left of orange arc).

REMARK 2.4. The motivation for defining the σ-atlas is to obtain statements on the singular spectrum of $aXY + bYX$ for prescribed singular values of X and Y. Since arbitrary arguments will be very hard to handle, we focus on those for which the non-zero values are all equal. The idea of constructing a skeleton made up of such slices is inspired by the fact that the singular values of a linear operator reflect the extremes of the unit ball's elliptical image and thus give a description of the mapping.

Our map could also be interpreted as if it was acting on vectors. But for these, spectral properties would be meaningless. Matrix arguments are naturally grouped by the rank. However, matrices of different rank can be arbitrary close to each other. Strictness regards only multiples of truncated isometries, for which the singular value vectors are somehow extremal. Note that we allow to extend matrices with zeros for the commutator, implying some independence of their order. In general, passing to larger sizes by such an embedding can possibly increase the regions $\mathcal{S}_{\tilde{r},r}(f)$.

3. A cartography of the commutator. Fig. 3.1 gives an impression of $\mathcal{S}_r$ with $r \neq 1$. The order of the matrices is taken to be r, which is the smallest possible value and motivated by the known equality cases of (1.2) (see [3, Proposition 4.6] for the so-called *maximal pairs*). Modified tests including larger matrices show no noteworthy change in the pictures. We seemingly obtain triangles, whose corners are not hard to guess. Once having an idea, verifying the assertion is straight forward.

The obtained images are misleading a bit. They suggest that the origin (at least for large r) does not belong to the set. This conclusion is completely false.

LEMMA 3.1. *$(0,0) \in \mathcal{S}_r$ for all $r \in \mathbb{N}$.*

Proof. We have to find commuting pairs subject to the constraints specified in the declaration of $\mathcal{S}_r$, resp. $\mathcal{S}_{1,r}(XY - YX)$. The two $(r+1) \times (r+1)$ matrices

$$X = I_1 \oplus O_r, \quad Y = \tfrac{1}{\sqrt{r}}(O_1 \oplus I_r)$$

meet these criteria. □

Knowing about convexity of the sets $\mathcal{S}_r$ or that they are star domains with center $(0,0)$ would allow us to concentrate solely on the boundaries. But, although $\mathcal{S}_r$ will turn out to satisfy both of these properties, we see no abstract way to a proof and must fill the gaps manually.

THEOREM 3.2. *Let $r \in \mathbb{N}$, $r \geq 2$. Then,*

$$\mathcal{S}_r = \operatorname{conv}\left\{(0,0),\ \left(\tfrac{2}{\sqrt{r}},0\right),\ \left(\tfrac{1}{\sqrt{r}},\tfrac{1}{\sqrt{r}}\right)\right\}$$

is the convex hull of three points.

Proof. In the same manner as (1.3),

$$\|XY - YX\|_{\mathsf{F}} \leq 2\|X\|_{\mathsf{F}}\|Y\|_{\infty} = \frac{2}{\sqrt{r}}.$$

This can be based on Ky Fan's inequalities by replacing the norm's submultiplicativity with the composite upper estimate $\|XY\|_{\mathsf{F}} \leq \|X\|_{\mathsf{F}}\|Y\|_{\infty}$. We also have the stronger estimate

$$\|XY - YX\|_1 \leq 2\|X\|_1\|Y\|_{\infty} = \frac{2}{\sqrt{r}},$$

whence $s_1 + s_2 \leq \frac{2}{\sqrt{r}}$. Since $\operatorname{rank} X = 1$, always $\|X\|_{\mathsf{F}} = \|X\|_1 = s_1(X)$. With the restriction to the first 45 degree octant, $\mathcal{S}_r$ is contained in the specified convex hull. For showing sharpness, we have to give examples with singular values on the outer border (red line in Fig. 3.1).

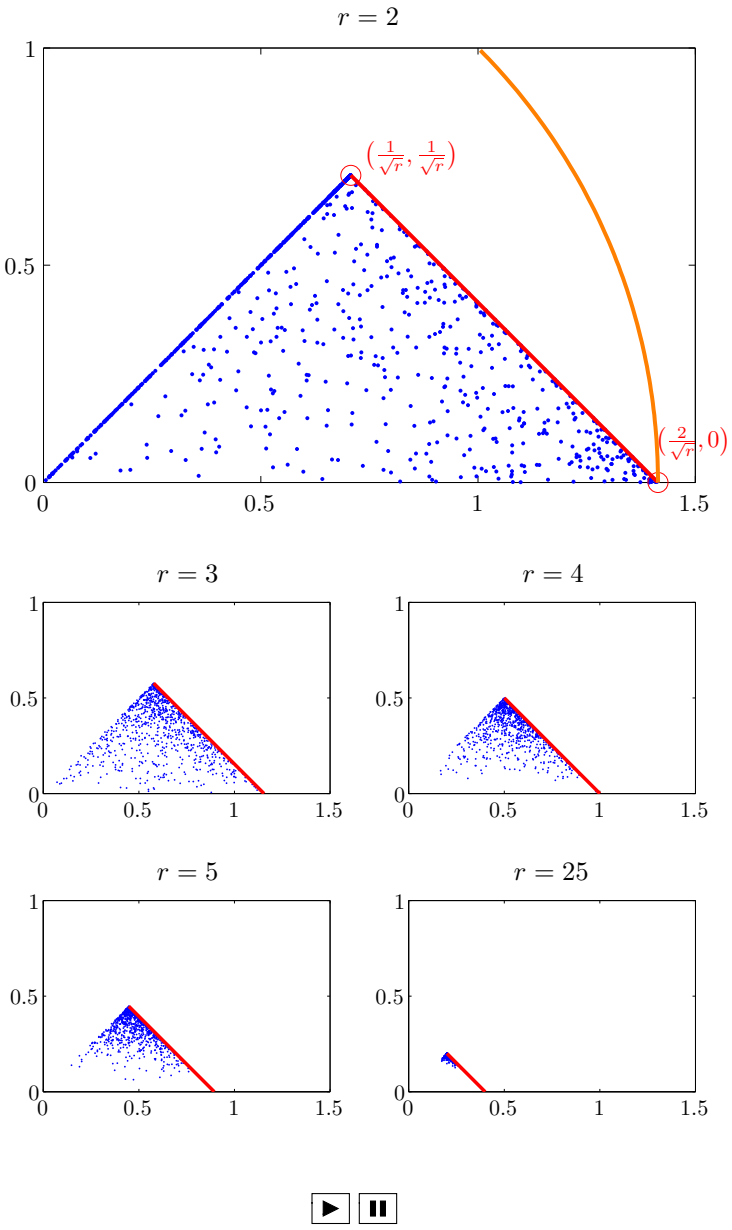


FIG. 3.1. *Appraising $\mathcal{S}_r$ for $r \geq 2$.*
We see singular values of the commutator for 1000 random matrix pairs $X \sim^{(1)} 1$, $Y \sim^{(r)} \frac{1}{\sqrt{r}}$ for different r. A video (or a couple of snapshots in the overview print) outlines the path to a rule for the line (red) that looks like the border. It does not coincide with the bound imposed by (1.2), depicted orange. The coordinates shown in the animation are rounded to two digits.

First, consider $r = 2$. Set

$$X_0 = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \quad Y_0 = \begin{bmatrix} b & a \\ a & -b \end{bmatrix}$$

with $a, b \in \mathbb{R}$. Since the columns of Y_0 are perpendicular, the matrix is 2-strict. Then, $X_0Y_0 - Y_0X_0 = \left[\begin{smallmatrix} a & -2b \\ 0 & -a \end{smallmatrix}\right]$ and, by (2.1), $s_{1,2}^2 = a^2 + 2b^2 \pm \sqrt{(a^2+2b^2)^2 - a^4}$, yielding

$$(s_1 + s_2)^2 = 2a^2 + 4b^2 + 2\sqrt{a^4} = 4a^2 + 4b^2.$$

If $a \in \left[0, \frac{1}{\sqrt{2}}\right]$ and $b = \sqrt{\frac{1-2a^2}{2}}$, one has $\|Y_0\|_{\mathsf{F}} = 1$ and gets $s_1 + s_2 = \sqrt{2}$, as desired.

Still, it needs to be clarified whether all points on the line may be obtained. Therefore, in the case $a = 0$, observe that $(s_1, s_2) = (\sqrt{2}, 0)$. When choosing $a = \frac{1}{\sqrt{2}}$, the result is $\left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right)$. The intermediate value theorem applied on s_1 (or s_2) now does the job.

So far, even for $r = 2$, it has been shown $\mathcal{S}_r \subset \operatorname{conv}\left\{(0,0),\ \left(\frac{2}{\sqrt{r}}, 0\right),\ \left(\frac{1}{\sqrt{r}}, \frac{1}{\sqrt{r}}\right)\right\}$, only. We appeal to specific constructions for proving the converse inclusion. For every choice of a (and b, accordingly), there is a rank one matrix commuting with Y_0; simply consider $X = \frac{1}{2}I_2 + \frac{1}{\sqrt{2}}Y_0$. Hence, with help of the intermediate value theorem, there is an X for every value between $s_1 + s_2 = 0$ and $s_1 + s_2 = \sqrt{2}$. Similar to our treatment of the outer boundary, the theorem can be utilised once more to prove existence of examples for every point on these lines.

In order to tackle the case $r > 2$, we require at least the matrix size $n = r$. With X_0, Y_0 as before, have a look at the pair

$$X = X_0 \oplus O_{r-2}, \quad Y = Y_0 \oplus \left(\tfrac{1}{\sqrt{2}} I_{r-2}\right).$$

Since $\sigma(Y_0) = \left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right)$, Y has strict rank r. The singular values of $XY - YX$ are the same as of $X_0Y_0 - Y_0X_0$. Certainly, Y furthermore has to been scaled with $\lambda = \sqrt{1 + (r-2)/2}$ for achieving norm 1. Thus, we have to multiply (s_1, s_2) from the previous case (for which the sum of the two coordinates is $\sqrt{2}$) by $\lambda^{-1} = \frac{\sqrt{2}}{\sqrt{r}}$, giving the asserted value $\frac{2}{\sqrt{r}}$. □

Up to now has been the easy part. Let us switch to the difficult case. For the remainder, X and Y are both rank one matrices with norm 1. The logical continuation of the last experiences would be to expect the singular values (s_1, s_2) inside the triangular set spanned by the points $(0,0)$, $(2,0)$, and $(1,1)$. Compared to the trivial estimate (1.3), this indeed would be an improvement. In fact, it is just the 1-norm-analogue of the ∞-norm-inequality. But, (1.2) actually delivers a better bound and entails that we should not rely on the triangle as the result can be at most a disc sector, which is a proper subset (see Fig. 3.2a).

When randomly selecting X, Y and portraying the corresponding singular value pairs, the picture is another one again (Fig. 3.2b) – also the smallest bound seems largely unreachable. We know that in principle all the points on the arc with radius $\sqrt{2}$ occur as singular value pairs. In contrast, Theorem 3.2 gives a growing distance of the singular value pairs to the arc if $\operatorname{rank} X = 1$ and Y is strict with rank $r \geq 2$. However, even some pictures in Fig. 3.1 have shown areas delusively left empty by random tests (compare with Lemma 3.1). But, as ranks are small and the distance to the circle is remarkably large, we tempted to believe that the blank region is not an issue of the experimental exploration. Let's gather some more information. The upcoming cases' numbers are coloured according to their illustration in Fig. 3.2c.

PROPOSITION 3.3. *The following statements restrict* $\mathcal{S}_1 \subset \mathbb{R}^2$.

a) *Except for* $(1,1)$ *the points in the plane with distance* $\sqrt{2}$ *to the origin do not belong to* $\mathcal{S}_1$.

b) *There are appropriate examples yielding* $(1,0) \in \mathcal{S}_1$. *Points further right on the* s_1*-axis cannot be attained.*

c) *All the points*

$$(s_1, s_2) = \frac{\sqrt{16 \cos\phi \, \sin\phi}}{1 + 2\, \cos\phi \, \sin\phi} \, (\cos\phi, \sin\phi) \tag{3.1}$$

with $\phi \in \left[0, \frac{\pi}{4}\right]$ *are in* $\mathcal{S}_1$.

d) *There exist rank one matrix pairs (both norm 1) whose commutator's singular values fill the line segment*

$$s_2 = s_1 - 1 \tag{3.2}$$

spanned by the points $(1,0)$ *and* $\left(\frac{\sqrt{2}+1}{2}, \frac{\sqrt{2}-1}{2}\right)$.

e) *The point* $\left(\frac{\sqrt{2}+1}{2}, \frac{\sqrt{2}-1}{2}\right)$ *is the unique intersection of the curves defined by the equations* (3.1) *and* (3.2).

Proof. a) In Section 3.3 of [9] it was noticed that there is no maximal pair of rank one matrices whose commutator also has rank one. Consequently, $(\sqrt{2}, 0) \notin \mathcal{S}_1$. The same argument applies to almost all points on the circle. By [3, Proposition 4.5], under these circumstances, without loss of generality $X = \mathbf{a}\mathbf{b}^*$, $Y = \mathbf{b}\mathbf{a}^*$ with unit vectors $\mathbf{a}, \mathbf{b}$ conforming to the orthogonality $\mathbf{a}^*\mathbf{b} = 0$. One easily observes

$$XY - YX = \mathbf{a}\mathbf{a}^* - \mathbf{b}\mathbf{b}^* = \begin{bmatrix} \mathbf{a} & \mathbf{b} \end{bmatrix} \begin{bmatrix} \mathbf{a}^* \\ -\mathbf{b}^* \end{bmatrix}$$

and $\sigma(XY - YX) = (1, 1, 0, \ldots)$. This example already appeared in [8, Proposition 6] in the context of (generalised) maximality.

b) For instance, $X = \left[\begin{smallmatrix} 0 & 1 \\ 0 & 0 \end{smallmatrix}\right]$ and $Y = \left[\begin{smallmatrix} 1 & 0 \\ 0 & 0 \end{smallmatrix}\right]$ yield $(1,0) \in \mathcal{S}_1$.

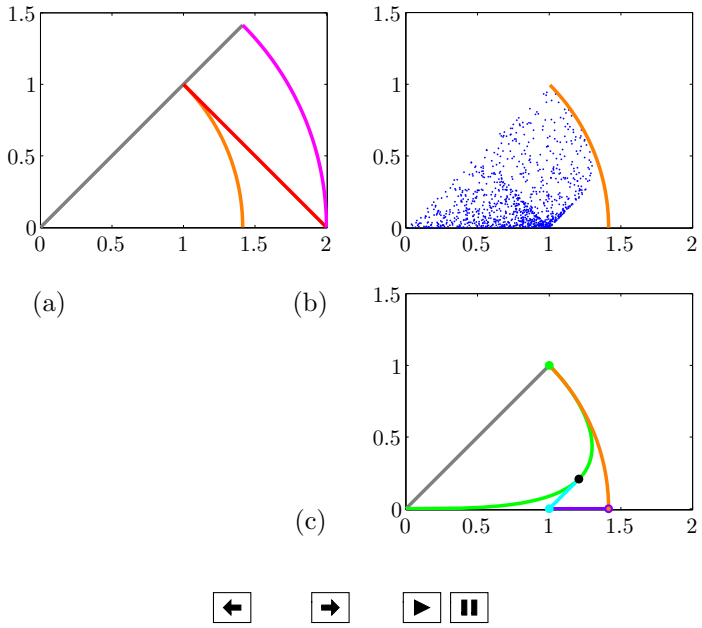


FIG. 3.2. *Transporting the result of the case $r \geq 2$ to a conjecture for $r = 1$ fails.*

(a) The obvious claim for bounding $\mathcal{S}_1$ based on Theorem 3.2 (red) lies between the trivial (1.3) (magenta) and sharp F-estimates (1.2) (orange).

(b) By testing 1000 pairs of normalised rank one matrices one obtains singular value pairs of their commutator (blue dots). The slideshow discusses the intricate result. A cryptic curve can be identified that fits the possible boundary of $\mathcal{S}_1$ quite good; details follow.

(c) There are points ensured to be not achievable by singular value pairs (orange and violet). That $\mathcal{S}_1$ contains the curves (3.1) (green) and (3.2) (cyan) is secured. They share a point of contact (black). Compare the image with (b) – the arrows will shift the sub-figure.

Now, assume $X = \mathbf{a}\mathbf{b}^*$, $Y = \mathbf{c}\mathbf{d}^*$, and

$$XY - YX = (\mathbf{c}, \mathbf{b})\mathbf{a}\mathbf{d}^* - (\mathbf{a}, \mathbf{d})\mathbf{c}\mathbf{b}^*$$

have rank one. Then, $(\mathbf{a}, \mathbf{d}) = 0$ or $\mathbf{b} = \kappa\mathbf{d}$ with $|\kappa| = 1$ (or similarly, $(\mathbf{c}, \mathbf{b}) = 0$ or $\mathbf{c} = \lambda\mathbf{a}$). When considering $(\mathbf{a}, \mathbf{d}) = 0$, the matrix $XY - YX = (\mathbf{c}, \mathbf{b})\mathbf{a}\mathbf{d}^*$ has norm

$$s_1(XY - YX) = \|(\mathbf{c}, \mathbf{b})\mathbf{a}\mathbf{d}^*\|_{\mathsf{F}} = |(\mathbf{c}, \mathbf{b})| \leq 1$$

as stated. If we suppose $\mathbf{b} = \kappa\mathbf{d}$ then $XY - YX = \overline{\kappa}[(\mathbf{c}, \mathbf{d})\mathbf{a} - (\mathbf{a}, \mathbf{d})\mathbf{c}]\mathbf{d}^*$. Decomposing $\mathbf{c} = \alpha\mathbf{a} + \beta\tilde{\mathbf{c}}$ with $\tilde{\mathbf{c}} \perp \mathbf{a}$ and $|\alpha|^2 + |\beta|^2 = 1$,

$$\|XY - YX\|_{\mathsf{F}} = |\beta|\|(\tilde{\mathbf{c}}, \mathbf{d})\mathbf{a} - (\mathbf{a}, \mathbf{d})\tilde{\mathbf{c}}\|,$$

whence $\mathbf{c} \perp \mathbf{a}$ may be demanded. Thus, analogously writing $\mathbf{d} = \alpha\mathbf{a} + \beta\mathbf{c} + \gamma\tilde{\mathbf{d}}$,

$$\|(\mathbf{c}, \mathbf{d})\mathbf{a} - (\mathbf{a}, \mathbf{d})\mathbf{c}\|^2 = |(\mathbf{c}, \mathbf{d})|^2 + |(\mathbf{a}, \mathbf{d})|^2 = |\alpha|^2 + |\beta|^2 \leq 1.$$

c) With (2.1) calculate the singular values of

$$X = \begin{bmatrix} s \\ c \end{bmatrix} \begin{bmatrix} c & -s \end{bmatrix} = \begin{bmatrix} cs & -s^2 \\ c^2 & -cs \end{bmatrix}, \quad Y = \begin{bmatrix} c \\ s \end{bmatrix} \begin{bmatrix} s & -c \end{bmatrix} = \begin{bmatrix} cs & -c^2 \\ s^2 & -cs \end{bmatrix},$$

where $s = \sin\phi$, $c = \cos\phi$.

d) Regarding the pairs

$$X = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \quad Y = \begin{bmatrix} c \\ -s \end{bmatrix} \begin{bmatrix} c & s \end{bmatrix} = \begin{bmatrix} c^2 & cs \\ -cs & -s^2 \end{bmatrix}$$

with $c = \cos\phi$, $s = \sin\phi$, (2.1) leads to

$$s_{1,2}^2 = \tfrac{1}{2} + t \pm \sqrt{\tfrac{1}{4} + t}, \quad \text{where } t = (\cos\phi \sin\phi)^2 \in \left[0, \tfrac{1}{4}\right].$$

We obtain $(1, 0)$ for $t = 0$. And for $t = \frac{1}{4}$, we get the point $\left(\frac{\sqrt{2}+1}{2}, \frac{\sqrt{2}-1}{2}\right)$. In between is a straight line since $\frac{ds_1}{dt} = \frac{ds_2}{dt}$. □

REMARK 3.4. The unforeseen curve appearing in Fig. 3.2 has been encountered before. It brings back into mind the curve (40) from [3] – well, seemingly multiplied with $\sqrt{2}$ in order to reverse the scaling done there. The result of a special group of examples was introduced earlier for a concrete purpose. One may consider inequalities like (1.2) with another norm, i.e.

$$\|XY - YX\|_\star \le C_\star \|X\|_\star \|Y\|_\star.$$

Of course, the number $C_\star$ on the right-hand side then can be different from $\sqrt{2}$ (e.g. 2 for $\|\cdot\|_\infty$ and $2^{\max\{1/p, 1-1/p\}}$ for the Schatten p norm). Under a natural normalisation assumption the constant cannot be lower than $\sqrt{2}$ (see [3, Proposition 5.1]). Fixing the size on 2×2 matrices, the Frobenius norm is, in fact, the only unitarily invariant one that allows this smallest constant in the estimate ([3, Theorem 5.4]). The proof was founded on a statement that with a point

$$\rho(\phi)(\cos\phi, \sin\phi)$$

also the point

$$\rho(\phi)^2 \left(\sqrt{\cos\phi \sin\phi} + \tfrac{\cos\phi - \sin\phi}{\sqrt{2}}, \sqrt{\cos\phi \sin\phi} - \tfrac{\cos\phi - \sin\phi}{\sqrt{2}}\right)$$

must be contained in the norm unit ball. The latter is the usual Euclidean unit disc for the Frobenius norm, and also the converse is true. The fascinating, unexpectable effect is that (3.1) (divided by $\sqrt{2}$) served as an excellent initial point for generating a sequence of curves converging to the unit circle (see [3, Remark 5.5]). Fig. 3.3 illuminates these proceedings. Note that for 3×3 matrices other norms than the Frobenius norm exist also having $\sqrt{2}$ as bounding constant (demonstrated in [4]).

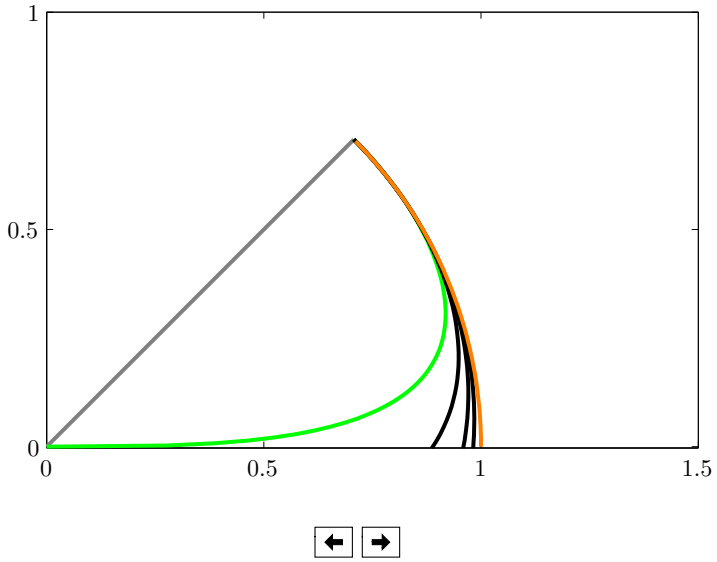


FIG. 3.3. *An old friend.*
Starting with a scaled by $1/\sqrt{2}$ version of (3.1) (green), an eighth of the unit circle (orange) is approximated. One point rises implications for another one. Although ρ^2 reduces the modulus (thus gains appearing unlikely), the smaller argument is able to compensate and results into monotone convergence of the sequence of curves (black). You can navigate through some steps of the iteration by clicking the arrows. The dependence between the curves' points is then also given a description.

REMARK 3.5. An incidental remark: the factor $\sqrt{2}$ in (1.2) suggests some kind of orthogonality between XY and YX. In fact, by

$$\|XY - YX\|_{\mathsf{F}}^2 = \|XY\|_{\mathsf{F}}^2 + \|YX\|_{\mathsf{F}}^2 - 2\,\mathrm{Re}(XY - YX) \leq 2$$

and

$$\|XY\|_{\mathsf{F}}, \|YX\|_{\mathsf{F}} \leq 1$$

one could think $\mathrm{Re}(XY - YX) \geq 0$. However, although in the real case [2, Remark 5.3] gives $\|XY - YX\|_{\mathsf{F}}^2 \leq \|XY\|_{\mathsf{F}}^2 + \|YX\|_{\mathsf{F}}^2$ whenever X commutes with Y^*, these ideas are not true in general. As we have seen the occurrence of X, Y, and $XY - YX$ being all rank one matrices, it can happen that XY and YX are linearly dependent. Commutativity is a trivial example. Besides this, both summands may have opposite directions, as demonstrated by

$$X = \tfrac{1}{\sqrt{2}}\left[\begin{smallmatrix} 1 & 0 \\ 0 & -1 \end{smallmatrix}\right], \quad Y = \tfrac{1}{\sqrt{2}}\left[\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right], \quad XY = \tfrac{1}{2}\left[\begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix}\right], \quad YX = \tfrac{1}{2}\left[\begin{smallmatrix} 0 & -1 \\ 1 & 0 \end{smallmatrix}\right]$$

known from maximality. But then, at least one of the two norms is far away from 1.

One might think that the curve (3.1) represents the outer boundary, but there are definitely some points in the outside. Fig. 3.2c strongly indicates an additional area bounded by a line segment. Moreover, we have just noticed that $\mathcal{S}_1$ does definitely not contain the whole disc sector. Instead, a more complicated boundary is indicated.

CONJECTURE 3.6. *$\mathcal{S}_1$ is the planar set enclosed by the curves* (3.1) *(taken for $s_2 \geq \frac{\sqrt{2}-1}{2}$),* (3.2) *(otherwise), and the rays $s_2 = 0$, $s_2 = s_1$ bounding the first octant.*

4. Treating the real case. The remainder is concerned with the verification of a special case of the conjecture.

THEOREM 4.1. *When restricted to real matrices, Conjecture 3.6 is true.*

The proof requires some work and basically consists of the following steps:

- Find a suitable description for the allowed matrices X and Y.
- Transform the singular value vectors of their commutator in such a way that the set's claimed boundaries have an accessible representation.
- Show several inequalities in order to bound $\mathcal{S}_1$ from above.
- Ensure the existence of an associated matrix pair for every point inside the proposed singular value set.

First, let $X = \mathbf{a}\mathbf{b}^*$, $Y = \mathbf{c}\mathbf{d}^*$, whence

$$XY - YX = \begin{bmatrix} \mathbf{a} & \mathbf{c} \end{bmatrix} \begin{bmatrix} (\mathbf{c},\mathbf{b}) & 0 \\ 0 & -(\mathbf{a},\mathbf{d}) \end{bmatrix} \begin{bmatrix} \mathbf{d}^* \\ \mathbf{b}^* \end{bmatrix}.$$

Here, $\mathbf{a},\mathbf{b},\mathbf{c},\mathbf{d} \in \mathbb{R}^n$ with $\|\mathbf{a}\| = \|\mathbf{b}\| = \|\mathbf{c}\| = \|\mathbf{d}\| = 1$ are considered. By appropriate unitary transformations from left and right,

$$\begin{aligned} XY - YX &\sim \begin{bmatrix} 1 & (\mathbf{c},\mathbf{a}) \\ 0 & \sqrt{1-|(\mathbf{c},\mathbf{a})|^2} \end{bmatrix} \begin{bmatrix} (\mathbf{c},\mathbf{b}) & 0 \\ 0 & -(\mathbf{a},\mathbf{d}) \end{bmatrix} \begin{bmatrix} 1 & 0 \\ (\mathbf{d},\mathbf{b}) & \sqrt{1-|(\mathbf{d},\mathbf{b})|^2} \end{bmatrix} \\ &= \begin{bmatrix} (\mathbf{c},\mathbf{b}) - (\mathbf{c},\mathbf{a})(\mathbf{a},\mathbf{d})(\mathbf{d},\mathbf{b}) & -(\mathbf{c},\mathbf{a})(\mathbf{a},\mathbf{d})\sqrt{1-|(\mathbf{d},\mathbf{b})|^2} \\ -(\mathbf{a},\mathbf{d})(\mathbf{d},\mathbf{b})\sqrt{1-|(\mathbf{c},\mathbf{a})|^2} & -(\mathbf{a},\mathbf{d})\sqrt{1-|(\mathbf{c},\mathbf{a})|^2}\sqrt{1-|(\mathbf{d},\mathbf{b})|^2} \end{bmatrix}, \end{aligned}$$

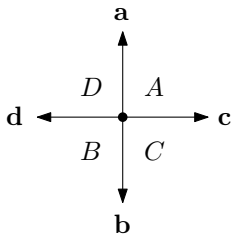
where only the non-trivial upper 2-by-2-block is taken into account. Singular values are not altered during this process. Abbreviating

$$A = (\mathbf{a},\mathbf{c}), \quad B = (\mathbf{b},\mathbf{d}), \quad C = (\mathbf{c},\mathbf{b}), \quad D = (\mathbf{d},\mathbf{a}),$$

we therefore get those of the commutator from the last matrix by employing (2.1):

$$s_{1,2}^2 = \tfrac{C^2+D^2}{2} - ABCD \pm \sqrt{\tfrac{(C^2-D^2)^2}{4} + (CD)^2(A^2+B^2) - (C^2+D^2)ABCD}. \tag{4.1}$$

Note that the transition to the real case cancels certain terms and allows to ignore several absolute values.



Since A, B, C, D are scalar products of vectors with norm one, all their values are between -1 and 1. Hence, $(2, 0)$ is in the range of the (4.1)-induced map

$$(A, B, C, D) \in [-1, 1]^4 \mapsto (s_1, s_2).$$

We know that this is not a point in $\mathcal{S}_1$. Indeed, we require $\mathbf{a} = \mathbf{d}$ and $\mathbf{a} = -\mathbf{d}$ to make that happen, which is impossible.

There are further restrictions to the domain generated by taking products of four vectors in a circle. Passing to angles ($A = \cos\alpha$ etc.) unveils these hidden links.

LEMMA 4.2. *Let $\mathbf{a}, \mathbf{b}, \mathbf{c}, \mathbf{d} \in \mathbb{R}^n$ be unit vectors. Suppose angles are named by*

$$\alpha = \angle(\mathbf{a}; \mathbf{c}), \quad \beta = \angle(\mathbf{b}; \mathbf{d}), \quad \gamma = \angle(\mathbf{b}; \mathbf{c}), \quad \delta = \angle(\mathbf{a}; \mathbf{d}).$$

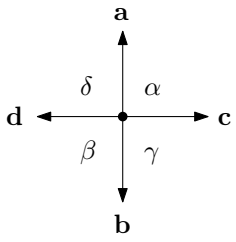
Then,

$$\begin{aligned}
\alpha &\le \beta + \gamma + \delta \le \alpha + 2\pi, \\
\beta &\le \alpha + \gamma + \delta \le \beta + 2\pi, \\
\gamma &\le \alpha + \beta + \delta \le \gamma + 2\pi, \\
\delta &\le \alpha + \beta + \gamma \le \delta + 2\pi.
\end{aligned}$$

Proof. The four left inequalities are simply seen similar to the triangle inequality (extended to more than three participants).

For showing the top right inequality, remember that three vectors are without loss of generality elements in $\mathbb{R}^3$. There exist unitary operations canceling the other entries. For such $\mathbf{a}, \mathbf{b}, \mathbf{c}$, clearly,

$$\alpha + \gamma + \angle(\mathbf{a}; \mathbf{b}) \le 2\pi.$$



The inequality is strict whenever $\alpha + \gamma < \pi$. Otherwise equality holds only if the vectors are coplanar.
Analogously to the left inequalities, we have

$$\beta \le \delta + \angle(\mathbf{a}; \mathbf{b})$$

and, consequently, $\alpha + \gamma + \beta - \delta \le 2\pi$ as stated.
The other three inequalities are obvious by swapping names. □

Fig. 4.1 gives a visualisation of these inequalities. Moreover, the additionally shown experimental data suggests that they are sharp. This gives rise to the hope that the fitting domain has been catched.

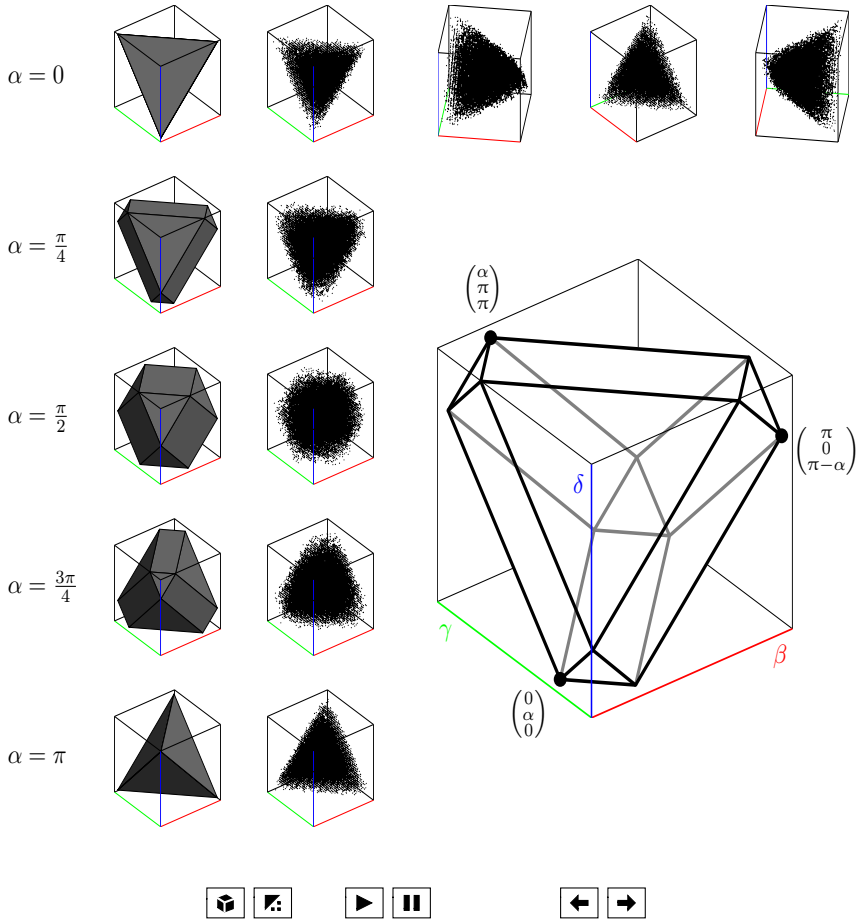


FIG. 4.1. *Restrictions of angles in four-vector-configurations.*

Lemma 4.2 given form: The 4-dimensional data can be made accessible by a series of slices that depend on a moving (but fixed at a time) α. The axes of the remaining variables β, γ, and δ are coloured red, green, and blue, resp., ranging roughly from 0 to π. The inequalities then can be interpreted as eight planes defining half-spaces. Combined with the trivial restrictions imposed by the cube $[0,\pi]^3$, we obtain the triple (β,γ,δ) to be inside a cubeoctahedron (or tetrahedron in degenerated cases).

Pushing the envelope: A test checks real vectors $\mathbf{a} = (1,0,0,0)$, $\mathbf{b} = (b_1,b_2,0,0)$, $\mathbf{c} = (c_1,c_2,c_3,0)$, $\mathbf{d} = (d_1,d_2,d_3,d_4)$ with norm 1 randomly. Thanks to isometry, this is sufficient. The attained quadruples $(\alpha,\beta,\gamma,\delta)$ are collected – rounded to a grid. Again, several 3D-cuts in dependence of one parameterised component α are shown. The raster plots resemble the exact cubeoctahedra quite impressively. You may toggle between precise results and experimental grid data with the second button.

Hints on the usage: The cube icon resumes to the interactive standard view. After activation, the scene can be rotated freely by dragging the mouse while holding down the left mouse button. Instead of an animated change playing back and forth in a loop (unavailable for pictures of the test), α can also be set to selected values via clicks on the arrows.

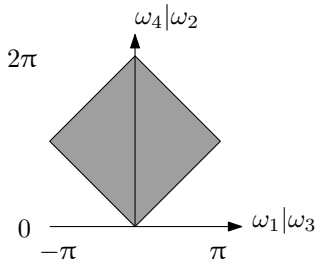
Handling the cyclically linked four angles is not easy. We decouple them and pass to two independent pairs in the following way:

$$\begin{aligned} \alpha - \beta &= \omega_1, & \gamma - \delta &= \omega_3 & (\omega_1, \omega_3 &\in [-\pi, \pi]), \\ \alpha + \beta &= \omega_2, & \gamma + \delta &= \omega_4 & (\omega_2, \omega_4 &\in [0, 2\pi]). \end{aligned}$$

The restrictions of the angles given in Lemma 4.2 now read

$$\begin{aligned} 0 \le -\omega_1 + \omega_4 \le 2\pi, \qquad & 0 \le \quad \omega_2 - \omega_3 \le 2\pi, \\ 0 \le \quad \omega_1 + \omega_4 \le 2\pi, \qquad & 0 \le \quad \omega_2 + \omega_3 \le 2\pi, \end{aligned}$$

and they can be easily given a graphical representation. The benefit is brought to light if one introduces the terms $W_i = \cos\omega_i$. In some sense, we return to the origins by transforming back to scalar products – only those related to other angles.



In the W-notation, the inequalities of Lemma 4.2 are equivalently restated by

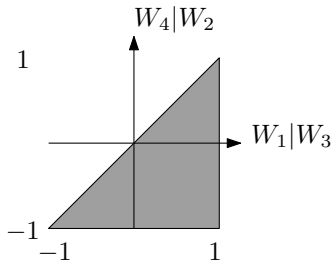
$$W_4 \le W_1 \quad \text{and} \quad W_2 \le W_3. \tag{4.2}$$

Eq. (4.1) can be rewritten as

$$4s_{1,2}^2 = -(W_1+W_2)(W_3+W_4) + 2(1+W_3W_4) \pm 2\sqrt{(1+W_1W_2)(W_3+W_4)^2 - (W_1+W_2)(W_3+W_4)(1+W_3W_4) + (1-W_3^2)(1-W_4^2)} \tag{4.3}$$

for describing the commutator's singular values.

The underlying domain of admissible points is converted from a set bounded by four lower and four upper inequalities to one which is determined by the inequality pair (4.2). The picture for the angle ω_i restrictions was a square. Their counterparts in scalar products W_i yield a half space. This is reduced to a triangle by the mapped cube bounds $-1 \le W_i \le 1$.



Based on these simple restrictions, it is not hard to convince oneself of $s_1 - s_2 \le 1$ by utilising standard methods of analysis. Calculating extremes of a real-valued function over interior points via partial derivatives and respecting points of non-differentiability yield that the set $\mathcal{S}_1$ is above the noticed line that seems to be half of its boundary. We are already pretty close to our aims and it ensures at least the existence of a large blank area inside the theoretically possible circular sector that vaguely has been predicted by the experiments near by $(\sqrt{2}, 0)$ (cf. Fig. 3.2b).

The next difficulty is the strange boundary of $\mathcal{S}_1$ composed of two curves which are positioned to each other such that a proper case differentiation appears barely possible – (3.1) bounds s_1 for large s_2, but for smaller ones, (3.2) is dominant instead. That's why we are going to substitute s_1 and s_2.

By (4.3), we know

$$s_1 = \frac{1}{2}\sqrt{x + \sqrt{x^2 - y^2}}, \quad s_2 = \frac{1}{2}\sqrt{x - \sqrt{x^2 - y^2}} \tag{4.4}$$

with the abbreviations

$$x = -(W_1 + W_2)(W_3 + W_4) + 2(1 + W_3W_4), \quad y = (W_1 - W_2)(W_3 + W_4).$$

It is not hard to verify

$$x = 2(s_1^2 + s_2^2), \quad |y| = 4s_1s_2, \tag{4.5}$$

a kind of inversion formula to (4.4). Actually, $(s_1, s_2) \leftrightarrow (x, y)$ may be interpreted as a coordinate transform. The big advantage in this point of view is demonstrated in Fig. 4.2: the two parts of $\mathcal{S}_1$'s apparent boundary now are the graph of a function, as it seems. Its law can be specified via direct calculation.

LEMMA 4.3. *In the x-y-system determined by* (4.5)*, the curve* (3.1) *is given by*

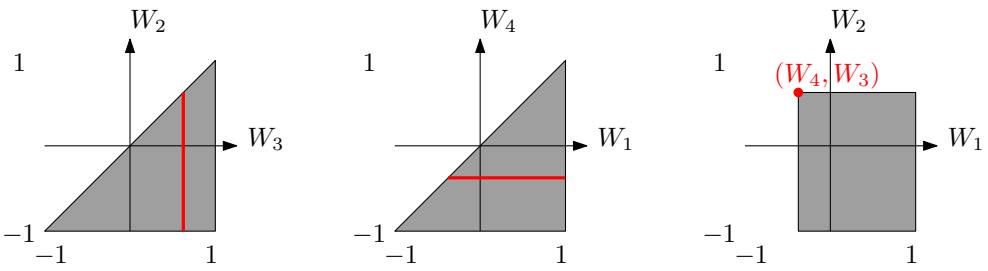
$$y = 8 - x - \sqrt{(8 - x)^2 - x^2},$$

the line (3.2) *follows the rule*

$$y = x - 2,$$

and their intersection $\left(\frac{\sqrt{2}+1}{2}, \frac{\sqrt{2}-1}{2}\right)$ *is mapped onto* $(3, 1)$*. The end points are transferred by* $(1, 0) \mapsto (2, 0)$ *and* $(1, 1) \mapsto (4, 4)$.

Finally, we translated our problem from $s_{1,2}(\mathbf{a}, \mathbf{b}, \mathbf{c}, \mathbf{d})$ into $\binom{x}{y}(W_1, W_2, W_3, W_4)$, in which a case differentiation can be properly done. To continue our plan, suppose W_3, W_4 are arbitrarily picked out of $[-1, 1]$, but fixed. Then, by (4.2), W_1, W_2 are allowed to be chosen from a specific interval only. This manifests in a vertical or a horizontal line in the domain picture. Consequently, (W_1, W_2) lies in the rectangle spanned by the points $(W_4, -1)$, $(1, -1)$, $(1, W_3)$, and (W_4, W_3):



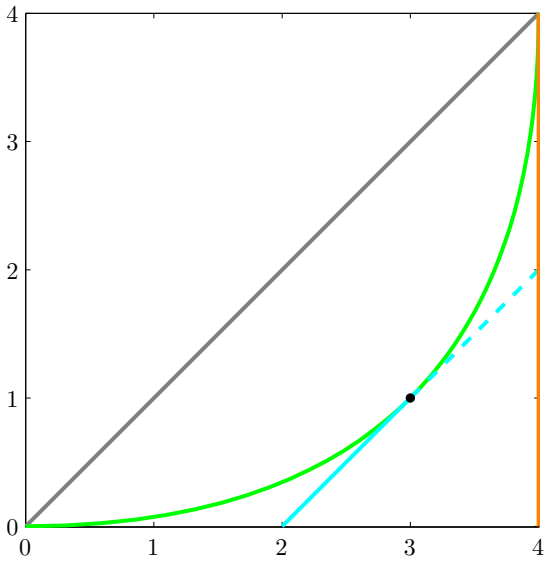


FIG. 4.2. *$\mathcal{S}_1$ in the x-y-system.*
With $(s_1, s_2) \mapsto (x, y)$ according to (4.5), the curves (3.1) and (3.2) (green and cyan), as well as their intersection (black point) become more handy. Further, the relation $s_1 - s_2 \leq 1$ is perfectly identifiable for the eye in the form $x - |y| \leq 2$. The $\sqrt{2}$-circle of the sharp F-estimate (1.2) becomes a vertical line in this system (orange).

The reduced function $(W_1, W_2) \mapsto (x, y)$ (with W_3, W_4 as parameters) is an affine linear mapping. Indeed, replacing $p = W_3 + W_4$ and $q = 2(1 + W_3 W_4)$,

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} -p & -p \\ p & -p \end{bmatrix} \begin{bmatrix} W_1 \\ W_2 \end{bmatrix} + \begin{bmatrix} q \\ 0 \end{bmatrix}.$$

Hence, it suffices to map the corners of the (W_1, W_2)-rectangle to determine its image. So, the two variables W_1, W_2 are eliminated. The four points $(W_4, -1)$, $(1, -1)$, $(1, W_3)$, and (W_4, W_3) correspond to

$$\begin{aligned}
P_{\text{I}} &= (2 + W_3 W_4 - W_4^2 + W_3 + W_4, && W_3 W_4 + W_4^2 + W_3 + W_4), \\
P_{\text{II}} &= (2 + 2W_3 W_4, && 2W_3 + 2W_4), \\
P_{\text{III}} &= (2 + W_3 W_4 - W_3 - W_4 - W_3^2, && W_3 + W_4 - W_3^2 - W_3 W_4), \\
P_{\text{IV}} &= (2 - W_3^2 - W_4^2, && W_4^2 - W_3^2).
\end{aligned}$$

The interior points of the original rectangle have images somewhere inside the quadrangle spanned by these four points. If $p = 0$, the latter collapse to a single point. When the sign of p changes, the order of the points in pictures is reversed.

Next, we show that all these points lie inside the asserted area, whence $\mathcal{S}_1$ cannot be larger. With help of Lemma 4.3, we may completely work in the x-y-system, allowing the use of more pleasant formulas. We drop the discussion on the sign of y, as it's not a really big issue. We have to regard only four cases, in each of which $W_3, W_4 \in [-1, 1]$ are chosen arbitrarily and independently of each other.

IV: Obviously, $x \le 2$. One also checks $|y| \le x$ and $|y| \le 2 - x$. The enclosed area suffices the claim of the theorem without any doubt.

III: By swapping $W_3 \leftrightarrow -W_4$ and multiplying y with -1, the cases I and III are the same. Hence, in both situations, one may check $y \le x$ and $-y \le x$. Further estimates will be postponed to I.

II: Again, validating $|y| \le x$ is no problem. Because of $|y| \le 4$ and

$$y^2 \;=\; 4W_3^2 + 4W_4^2 + 8W_3W_4 \;\ge\; 16W_3W_4 \;=\; 8(x-2),$$

we also obtain $x - 2 \le \frac{y^2}{8} \le |y|$. By the way, this is equivalent to $s_1 - s_2 \le 1$. Now, we derive benefit from the system's representation for the most interesting part. The chain (of rather raw estimates)

$$|y| \;\ge\; \sqrt{8(x-2)} \;\ge\; 2x - 4 \;\ge\; 3x - 8 \;\ge\; 8 - x - \sqrt{(8-x)^2 - x^2}$$

is valid for $3 \le x \le 4$. Clearly, values $x > 4$ may not be achieved.

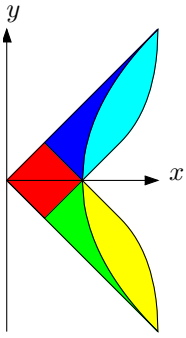
I: In the current situation, also $y \ge x - 2$ and $x \le 4$ are easily verified. The leftover is the challenging part of the boundary associated with $3 \le x \le 4$:

$$y \;=\; x - 2 + 2W_4^2 \;\ge\; 8 - x - \sqrt{64 - 16x} \tag{4.6}$$

is true if and only if x is between the roots $3 - W_4^2 \pm 2|W_4|$, i.e.

$$3 - W_4^2 - 2|W_4| \;\le\; 2 + W_3W_4 - W_4^2 + W_3 + W_4 \;\le\; 3 - W_4^2 + 2|W_4|. \tag{4.7}$$

The second inequality is always true, and the first one as long as $x \ge 3$.



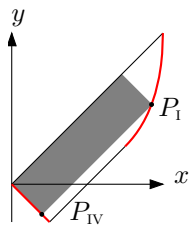
The provided estimates are of different strength. The picture on the left side highlights the possible locations of the four corners:

P_{I} P_{II} P_{III} P_{IV}

The red square is folded to a triangle when passing to absolute values. This element even manifests in the outcome of Fig. 3.2b by a visibly higher point density; it is hit more often in tests.

The very last step is showing that every point of the described set has to be in $\mathcal{S}_1$. We act in a similar fashion as we have done in the proof of Theorem 3.2.

By (3.1) and (3.2), each point on the right curve limiting the cyan and yellow areas is connected to certain (W_1, W_2, W_3, W_4). All of them are a corner P_{I} or P_{III} of a quadrangle that must be part of $\mathcal{S}_1$ (revealing the advantage of the reduction to extremal points of convex sets). We need to show that the union covers everything.



Analysing the equality cases of (4.6), we are required to look at the right end point of the interval which restricts x; see (4.7). These thoughts lead to $W_3 = 1$ and $W_4 \geq 0$. In fact, for $W_4 = 0$, we have $P_{\text{I}} = (3, 1)$, and we arrive at $P_{\text{I}} = (4, 4)$ with $W_4 = 1$. In between, we are passing through the outer boundary. Similarly, P_{IV} is on the path spanned by $(1, -1)$ and $(0, 0)$. All the created lines have slope 1 due to $P_{\text{I}} - P_{\text{IV}} = (2 + 2W_4, 2 + 2W_4)$. This finishes the proof of Theorem 4.1. □

In the words of the referee: It is a shame that the complex case isn't yet understood. But, the ideas used in the proof of the shape relied on the real numbers (most prominently, the concept of angles). In addition, Conjecture 3.6 is not a mere consequence of the real result. Nevertheless, computer experiments don't display any differences between real and complex inputs.

The determined sets $\mathcal{S}_r$ correspond to one row of the commutator's σ-atlas. Although the pattern's anomaly is interesting on its own, this gives a (partial) foundation of the overwhelming presence of almost commuting matrix pairs. The rank has much more freedom with increasing matrix size, and the majority of the pairs yields small singular values. Apart from determining the other sets $\mathcal{S}_{\tilde{r},r}(XY - YX)$, it would be interesting whether one can find an interpolation property to characterise the singular values of the commutator of arbitrary matrices.

REFERENCES

[1] R. Bhatia. *Matrix Analysis*, Springer-Verlag, New York, 1997.
[2] A. Böttcher, D. Wenzel. How big can the commutator of two matrices be and how big is it typically? *Linear Algebra Appl.*, 403:216–228, 2005.
[3] A. Böttcher, D. Wenzel. The Frobenius norm and the commutator. *Linear Algebra Appl.*, 429:1864–1885, 2008.
[4] K.-S. Fong, C.-M. Cheng, I.-K. Lok. Another unitarily invariant norm attaining the minimum norm bound for commutators. *Linear Algebra Appl.*, 433:1793–1797, 2010.
[5] R.A. Horn, C.R. Johnson. *Matrix Analysis*, Cambridge University Press, Cambridge, 1985.
[6] Z. Lu. Proof of the normal scalar curvature conjecture. arXiv:0711.3510v1 [math.DG], November 22 2007.
[7] S.-W. Vong, X.-Q. Jin. Proof of Böttcher and Wenzel's conjecture. *Oper. Matrices*, 2:435–442, 2008.
[8] D. Wenzel. Dominating the commutator. *Oper. Theory: Adv. and Appl.*, 202:579–600, 2010.
[9] D. Wenzel, K.M.R. Audenaert. Impressions of convexity – An illustration for commutator bounds. *Linear Algebra Appl.*, 433:1726–1759, 2010.