A Note on von Neumann's Trace Inequality

By ROLF DIETER GRIGORIEFF*

The following trace inequality first proved by von Neumann [6] has found repeated interest in the literature:

Theorem. *Let A,B be complex n×n matrices. Then*

$$(*) \qquad |\,\mathrm{tr}(AB)\,| \le (\sigma(A),\, \sigma(B))$$

where $\sigma(A) = (\sigma_1,\sigma_2,...,\sigma_n)$ *denotes the ordered vector of singular values* $\sigma_1 \ge \sigma_2 \ge ... \ge \sigma_n$ *of A and* $(\,\cdot\,,\cdot\,)$ *the euclidean scalar product.*

The importance of this result stems from the fact that it is used as an essential tool in Schatten's theory of cross-spaces and more recently in Ball's treatment of the equations of nonlinear elasticity.

The original proof was not particularly easy. While Schatten's representation essentially followed the original one (see e.g. [7, p. 58]) it was Mirsky [4,5] who presented two new ways in obtaining the result. The second one in [5] relies on a property of double stochastic matrices which is easy enough to access. Ball [1] and Ciarlet [2, p. 99] adapted this version in their representation. Nevertheless, Ciarlet remarks that 'unexpectedly, finding a decent proof of this seemingly simple result turns out to be anything but trivial'. The aim of the present note is to present still a further proof which seems to be elementary enough to correspond to the simplicity of the statement in von Neumann's Theorem.

P r o o f . It is a basic fact that *A,B* have the canonical representation (see e.g. [3, p. 261])

$$A = \sum_{k=1}^{n} \sigma_k(A)\,(\,\cdot\,\,, \varphi_k)\,\varphi_k' \,, \qquad B = \sum_{k=1}^{n} \sigma_k(B)\,(\,\cdot\,\,, \psi_k)\,\psi_k'$$

* Fachbereich Mathematik, Technische Universität Berlin, Strasse d. 17. Juni 135, 10623 Berlin, Germany.
e-mail: grigo@math.tu-berlin.de

where (φ_k), (φ_k'), (ψ_k), (ψ_k') are orthonormal families in $\mathbb{C}^n$. Also, for any orthonormal basis (χ_k)

$$\mathrm{tr}(AB) = \sum_{k=1}^{n} (AB\chi_k, \chi_k) \ .$$

Since both sides of $(*)$ are positive homogeneous with respect to $\sigma(A)$ resp. $\sigma(B)$ it is sufficient to consider the case $\sigma(A)_1 \le 1$, $\sigma(B)_1 \le 1$ only. Moreover, the left-hand side of $(*)$ is a convex function of $\sigma(A)$ resp. $\sigma(B)$ while the right-hand side is linear. Hence, each σ lying in the convex hull of the vectors $e_k := (1,...,1,0,...,0)$ with k components equal to 1, $k = 0,...,n$, we need to prove $(*)$ for the case $\sigma(A) = e_r$, $\sigma(B) = e_s$, $r,s = 0,...,n$, only. Due to $\mathrm{tr}(AB) = \mathrm{tr}(BA)$ it is sufficient to consider the case $r \le s$. Choosing $\chi_k = \psi_k$ we then obtain

$$|\,\mathrm{tr}(AB)\,| = |\sum_{j=1}^{n} \sum_{k=1}^{r} \sum_{\ell=1}^{s} (\psi_\ell', \varphi_k)\ (\chi_j, \psi_\ell)\ (\varphi_k', \chi_j)\ |$$

$$= |\sum_{k=1}^{r} \sum_{\ell=1}^{s} (\psi_\ell', \varphi_k)\ (\varphi_k', \psi_\ell)\ | \le \sum_{k=1}^{r} \|\,\varphi_k\,\|\,\|\,\varphi_k'\,\| = r \ .$$

But also $(\sigma(A), \sigma(B)) = r$ in this case. ■

Remark. Let $A \in \mathcal{B}_2(H, H')$, $B \in \mathcal{B}_2(H', H)$, the Schmidt class of operators between separable Hilbert spaces H and H' (see e.g. [3]). For these operators the ordered vector of singular value is defined as an element in the sequence space ℓ^2 and the theorem still holds true.

The proof follows easily from the fact that the canonical representation and the trace definition remain valid for $A \in \mathcal{B}_2(H, H')$, $B \in \mathcal{B}_2(H', H)$ when substituting n by ∞ and that the truncated series of A,B are convergent to A resp. B.

Remark (cp. [2, p. 280]). Let $\alpha,\beta \in \ell^2$ such that $|\alpha_1| \ge |\alpha_2| \ge ...$, $|\beta_1| \ge |\beta_2| \ge ...$. Then for any injective mapping $\pi : \mathbb{N} \to \mathbb{N}$

$$|\sum_{k=1}^{\infty} \alpha_k \beta_{\pi(k)}\,| \le \sum_{k=1}^{\infty} |\,\alpha_k \beta_k\,| \ .$$

The proof is obtained by applying the theorem to diagonal A,B.

References

[1] J.M. BALL, Convexity condition and existence theorems in nonlinear elasticity, Arch. Rat. Mech. Anal. **63** (1977) 337–403

[2] P.G. CIARLET, Mathematical Elasticity, Vol. I, North-Holland, Amsterdam, 1988

[3] T. KATO, Perturbation theory of linear operators, Springer, New York, 1966

[4] L. MIRSKY, On the trace of matrix products, Math. Nachr. **20** (1959) 171–174

[5] –, A trace inequality of John von Neumann. Monatsh. für Math. **79** (1975) 303-306

[6] J. VON NEUMANN, Some matrix-inequalities and metrization of matric-space, Tomsk Univ. Rev. **1** (1937) 286-300

[7] R. SCHATTEN, Norm ideals of completely continuous operators, Springer, Berlin, 1970