

## NORM ESTIMATES FOR FUNCTIONS OF TWO COMMUTING MATRICES\*

MICHAEL GIL<sup>†</sup>

**Abstract.** Matrix valued analytic functions of two commuting matrices are considered. A precise norm estimate is established. As a particular case, the matrix valued functions of two matrices on tensor products of Euclidean spaces are explored.

**Key words.** Functions of commuting matrices, Norm estimate.

**AMS subject classifications.** 15A15, 15A54, 15A69, 15A60.

**1. Introduction and statement of the main result.** In the book [4], I.M. Gel'fand and G.E. Shilov have established an estimate for the norm of a regular matrix-valued function in connection with their investigations of partial differential equations. However that estimate is not sharp, it is not attained for any matrix. The problem of obtaining a precise estimate for the norm of a matrix-valued function has been repeatedly discussed in the literature. In the paper [5] (see also [7]) the author has derived a precise estimate for a regular matrix-valued function. It is attained in the case of normal matrices. In the present paper we generalize the main result of the paper [5] to functions of two commuting matrices. Besides, the main result of the present paper-Theorem 1.1 is improved in the case of matrices on tensor products of Euclidean spaces.

It should be noted that functions of commuting operators were investigated by many mathematicians, cf. [1, 10, 12] and references therein however the norm estimates were not considered, but as it is well-known, matrix valued functions are Green's functions and characteristic functions of various differential and difference equations. This fact allow us to investigate stability, well-posedness and perturbations of these equations by norm estimates for matrix valued functions, cf. [2, 3, 6].

Let  $\mathbb{C}^n$  be a Euclidean space with a scalar product  $(\cdot, \cdot)$ , the unit matrix  $I$  and the Euclidean norm  $\|\cdot\| = (\cdot, \cdot)^{1/2}$ ;  $M(\mathbb{C}^n)$  denotes the set of all linear operators in  $\mathbb{C}^n$ . For a  $A \in M(\mathbb{C}^n)$ ,  $\|A\|$  is the operator norm;  $N(A)$  is the Frobenius (Hilbert-Schmidt) norm:  $N^2(A) = \text{Trace}(AA^*)$ ;  $A^*$  is the operator adjoint to  $A$ ,  $\lambda_j(A)$ ,  $j = 1, \dots, n$  are the eigenvalues counting with their multiplicities,  $\sigma(A)$  is the spectrum.

Everywhere below  $A$  and  $B$  are commuting matrices. Let  $\Omega_A$  and  $\Omega_B$  be open simple connected sets containing  $\sigma(A)$  and  $\sigma(B)$ , respectively. Let  $f$  be a scalar function analytic on  $\Omega_A \times \Omega_B$ . Introduce the operator valued function

$$(1.1) \quad f(A, B) := -\frac{1}{4\pi^2} \int_{L_B} \int_{L_A} f(z, w) R_z(A) R_w(B) dw dz,$$

---

\*Received by the editors 26 September 2004. Accepted for publication 6 March 2005. Handling Editor: Harm Bart.

<sup>†</sup>Department of Mathematics, Ben Gurion University of the Negev, Beer-Sheva 84105, Israel (gilmi@cs.bgu.ac.il).

where  $L_A \subset \Omega_A, L_B \subset \Omega_B$  are closed contour surrounding  $\sigma(A)$  and  $\sigma(B)$ , respectively. Note that if the series

$$f(z, w) = \sum_{j,k=0}^{\infty} c_{jk} z^j w^k$$

converges for  $|z| \leq r_s(A), |w| \leq r_s(w)$ , where  $r_s(A)$  denotes the spectral radius of  $A$ , then (1.1) holds.

The following quantity plays a key role in the sequel

$$g(A) = (N^2(A) - \sum_{k=1}^n |\lambda_k(A)|^2)^{1/2}.$$

Since

$$\sum_{k=1}^n |\lambda_k(A)|^2 \geq |\text{Trace } A^2|, \text{ we have } g^2(A) \leq N^2(A) - |\text{Trace } A^2|.$$

If  $A$  is a normal matrix, i.e. if  $AA^* = A^*A$ , then  $g(A) = 0$ . Also the inequality

$$(1.2) \quad g^2(A) \leq \frac{1}{2} N^2(A^* - A)$$

is valid, cf. [7, Section 2.1]. Introduce the numbers

$$\eta_k := \frac{1}{k!} \sqrt{\frac{(n-1)!}{(n-k-1)!(n-1)^k}} \text{ for } k = 1, \dots, n-1; \eta_0 = 1.$$

It is simple to check that

$$(1.3) \quad \eta_k^2 \leq \frac{1}{(k!)^3} \quad (k = 1, \dots, n-1).$$

Denote by  $co(A), co(B)$  the closed convex hulls of  $\sigma(A)$  and  $\sigma(B)$ , respectively. Put

$$f^{(j,k)}(z, w) = \frac{\partial^{j+k} f(z, w)}{\partial z^j \partial w^k}.$$

Now we are in a position to formulate the main result of the paper.

**THEOREM 1.1.** *Let  $A$  and  $B$  be commuting  $n \times n$ -matrices and  $f(z, w)$  be regular on a neighborhood of  $co(A) \times co(B)$ . Then*

$$\|f(A, B)\| \leq \sum_{j,k=0}^{j+k \leq n-1} \eta_j \eta_k g^j(A) g^k(B) \sup_{z \in co(A), w \in co(B)} |f^{(j,k)}(z, w)|.$$

The proof of this theorem is divided into a series of lemmas which are presented in the next section. If both  $A$  and  $B$  are normal operators, and

$$\sup_{z \in co(A), w \in co(B)} |f(z, w)| = \sup_{z \in \sigma(A), w \in \sigma(B)} |f(z, w)|,$$

then Theorem 1.1 gives us the exact relation

$$\|f(A, B)\| = \sup_{z \in \sigma(A), w \in \sigma(B)} |f(z, w)|.$$

Taking into account (1.3), we get

COROLLARY 1.2. *Under the hypothesis of Theorem 1.1, the estimate*

$$\|f(A, B)\| \leq \sum_{j,k=0}^{j+k \leq n-1} \frac{g^j(A)g^k(B)}{(j!k!)^{3/2}} \sup_{z \in co(A), w \in co(B)} |f^{(j,k)}(z, w)|$$

is true.

Let  $A$  be a normal matrix. Then  $g(A) = 0$ . Now Corollary 1.2 implies

$$\|f(A, B)\| \leq \sum_{0 \leq k \leq n-1} \frac{g^k(B)}{(k!)^{3/2}} \sup_{z \in co(A), w \in co(B)} \left| \frac{\partial^k f(z, w)}{\partial w^k} \right|.$$

Let us evaluate the error of Theorem 1.1 in the case of non-normal matrices. Certainly, we can obtain the exact value of the norm of a function of two matrices in very simple cases only. Consider the  $2 \times 2$ -matrices

$$A = \begin{pmatrix} a & 1/3 \\ 0 & a \end{pmatrix}.$$

and  $B = 2A$  with  $a > 0$ . Construct a function of two variables by setting  $(x, y) \rightarrow f(x + y)$ . Direct calculations show that

$$f(A + B) = \begin{pmatrix} f(3a) & f'(3a) \\ 0 & f(3a) \end{pmatrix}.$$

Assume that  $f(3a) = 0$ . Then  $\|f(A + B)\| = |f'(3a)|$ . At the same time, Theorem 1.1 gives us the relations,

$$\|f(A + B)\| \leq |f(a)| + |f'(3a)|(g(A) + g(B)) = |f'(3a)|,$$

since  $g(A) = 1/3$ ,  $g(B) = 2/3$ . Thus in the considered case Theorem 1.1 gives us the exact result.

EXAMPLE 1.3. Consider the polynomial

$$P(z, w) = \sum_{\nu=0}^{m_1} \sum_{l=0}^{m_2} c_{\nu l} z^\nu w^l$$

with complex, in general, coefficients. Then

$$P^{(j,k)}(z, w) = \sum_{\nu=0}^{m_1-j} \sum_{l=0}^{m_2-k} c_{\nu l} \nu(\nu-1)\dots(\nu-j+1) z^{\nu-j} l\dots(l-k+1) w^{l-k}.$$

Now Theorem 1.1 implies

$$\|P(A, B)\| \leq \sum_{j,k=0}^{j+k \leq n-1} \eta_j \eta_k g^j(A) g^k(B) \sum_{\nu=0}^{m_1-j} \sum_{l=0}^{m_2-k} \frac{\nu! l! |c_{\nu l}|}{(l-k)! (\nu-j)!} r_s^{\nu-j}(A) r_s^{l-k}(B).$$

Recall that  $r_s(.,.)$  denotes the spectral radius. If both  $A$  and  $B$  are normal operators, then

$$\|P(A, B)\| \leq \sum_{\nu=0}^{m_1} \sum_{l=0}^{m_2} |c_{\nu l}| r_s^{\nu}(A) r_s^l(B).$$

EXAMPLE 1.4. Consider the function

$$f(z, w) = \cos(xz + yw) \quad (y, x \geq 0).$$

Note that the function  $U(x, y) = \cos(xA + yB)$  is a solution of the equation

$$\partial^2 U(x, y) / \partial x^2 + \partial^2 U(x, y) / \partial y^2 + (A^2 + B^2)U(x, y) = 0.$$

In the considered case

$$|f^{(j,k)}(z, w)| = x^j y^k |\cos(xz + yw)| \quad (j+k \text{ is even}),$$

$$|f^{(j,k)}(z, w)| = x^j y^k |\sin(xz + yw)| \quad (j+k \text{ is odd}).$$

For simplicity assume that the spectra of both  $A$  and  $B$  are real. Then thanks to Theorem 1.1,

$$\|U(x, y)\| \leq \sum_{j,k=0}^{j+k \leq n-1} \eta_j \eta_k g^j(A) g^k(B) x^j y^k \quad (x, y \geq 0).$$

**2. Proof of Theorem 1.1.** The following lemma is needed.

LEMMA 2.1. Let  $\Omega$  and  $\tilde{\Omega}$  be the closed convex hulls of complex, in general, points

$$(2.1) \quad x_0, x_1, \dots, x_n$$

and

$$(2.2) \quad y_0, y_1, \dots, y_m,$$

respectively, and let a scalar-valued function  $f(z, w)$  be regular on  $D \times \tilde{D}$ , where  $D$  and  $\tilde{D}$  are neighborhoods of  $\Omega$  and  $\tilde{\Omega}$ , respectively. In addition, let  $L \subset D, \tilde{L} \subset \tilde{D}$  be Jordan closed contours surrounding the points in (2.1) and (2.2), respectively. Then with the notation

$$Y(x_0, \dots, x_n; y_0, \dots, y_m) = -\frac{1}{4\pi^2} \int_L \int_{\tilde{L}} \frac{f(z, w) dz dw}{(z - x_0) \cdots (z - x_n)(w - y_0) \cdots (w - y_m)},$$

we have

$$|Y(x_0, \dots, x_n; y_0, \dots, y_m)| \leq \frac{1}{n!m!} \sup_{z \in \Omega, w \in \tilde{\Omega}} |f^{(n,m)}(z, w)|.$$

*Proof.* First, let all the points be distinct:  $x_j \neq x_k$ ,  $y_j \neq y_k$  for  $j \neq k$ . Let a function  $h$  of one variable be regular on  $D$  and  $[x_0, x_1, \dots, x_n]h$  be a divided difference of function  $h$  at points  $x_0, x_1, \dots, x_n$ . Then

$$(2.3) \quad [x_0, x_1, \dots, x_n]h = \frac{1}{2\pi i} \int_L \frac{h(\lambda) d\lambda}{(\lambda - x_0) \cdots (\lambda - x_n)}$$

(see [4, formula (54)]). Thus

$$Y(x_0, \dots, x_n; y_0, \dots, y_m) = \frac{1}{2\pi i} \int_{\tilde{L}} \frac{[x_0, \dots, x_n]f(\cdot, w) dw}{(w - y_0) \cdots (w - y_m)}.$$

Now apply (2.3) to  $[x_0, \dots, x_n]f(\cdot, w)$ . Then

$$Y(x_0, \dots, x_n; y_0, \dots, y_m) = [x_0, \dots, x_n] [y_0, \dots, y_m] f(\cdot, \cdot) \equiv [x_0, \dots, x_n] ([y_0, \dots, y_m] f(\cdot, \cdot)).$$

The following estimate is well-known [11, p. 6]:

$$|[x_0, \dots, x_n] [y_0, \dots, y_m] f(\cdot, \cdot)| \leq \sup_{z \in \Omega, w \in \tilde{\Omega}} |f^{(n,m)}(z, w)|.$$

It proves the required result if all the points are distinct. If some points coincide, then the claimed inequality can be obtained by small perturbations and the previous arguments.  $\square$

Furthermore, since  $A$  and  $B$  commute they have the same orthogonal normal basis of the triangular representation (Schur's basis)  $\{e_k\}$ . We can write

$$(2.4) \quad A = D_A + V_A, B = D_B + V_B,$$

where  $D_A, D_B$  are the diagonal parts,  $V_A$  and  $V_B$  are the nilpotent parts of  $A$  and  $B$ , respectively. Furthermore, let  $|V_A|$  be the operator whose matrix elements in  $\{e_k\}$  are the absolute values of the matrix elements of the nilpotent part  $V_A$  of  $A$  with respect to this basis. That is,

$$|V_A| = \sum_{k=1}^n \sum_{j=1}^{k-1} |a_{jk}| (\cdot, e_k) e_j,$$

where  $a_{jk} = (Ae_k, e_j)$ . Similarly  $|V_B|$  is defined.

LEMMA 2.2. Under the hypothesis of Theorem 1.1 the estimate

$$\|f(A, B)\| \leq \sum_{j,k=0}^{j+k \leq n-1} \sup_{z \in co(A), w \in co(B)} |f^{(j,k)}(z, w)| \frac{\| |V_A|^j |V_B|^k \|}{j!k!}$$

is true, where  $V_A$  and  $V_B$  are the nilpotent part of  $A$  and  $B$ , respectively.

*Proof.* It is not hard to see that the representation (2.4) implies the equality

$$(A - I\lambda)^{-1} = (D_A + V_A - \lambda I)^{-1} = (I + R_\lambda(D_A)V_A)^{-1}R_\lambda(D_A)$$

for all regular  $\lambda$ . According to Lemma 1.7.1 from [7]  $R_\lambda(D_A)V_A$  is a nilpotent operator because  $V$  and  $R_\lambda(D_A)$  the same invariant subspaces. Hence,  $(R_\lambda(D_A)V_A)^n = 0$ . Therefore,

$$R_z(A) = \sum_{k=0}^{n-1} (R_z(D_A)V_A)^k (-1)^k R_z(D_A).$$

Similarly,

$$R_\mu(B) = \sum_{k=0}^{n-1} (R_\mu(D_B)V_B)^k (-1)^k R_\mu(D_B).$$

So from (1.1) it follows

$$(2.5) \quad f(A, B) = \sum_{j,k=0}^{n-1} C_{jk},$$

where

$$C_{jk} = \frac{(-1)^{k+j}}{4\pi^2} \int_{L_B} \int_{L_A} f(z, w) (R_z(D_A)V_A)^j R_z(D_A) (R_w(D_B)V_B)^k R_w(D_B) dz dw.$$

Since  $D_A$  is a diagonal matrix with respect to the Schur basis  $\{e_k\}$  and its diagonal entries are the eigenvalues of  $A$ , then

$$R_z(D_A) = \sum_{j=1}^n \frac{Q_j}{\lambda_j(A) - z},$$

where  $Q_k = (., e_k)e_k$ . Similarly,

$$R_z(D_B) = \sum_{j=1}^n \frac{Q_j}{\lambda_j(B) - z}.$$

Taking into account that  $Q_s V_A Q_m = 0$ ,  $Q_s V_B Q_m = 0$  ( $s \geq m$ ), we get

$$C_{jk} = \sum_{1 \leq s_1 < s_2 < \dots < s_{j+1} < m_1 < m_2 < \dots < m_{k+1} < n} Q_{s_1} V_A Q_{s_2} V_A \dots$$

$$V_A Q_{s_{j+1}} Q_{m_1} V_B Q_{m_2} V_B \dots V_B Q_{m_{k+1}} J(s_1, \dots, s_{j+1}, m_1, \dots, m_{k+1}),$$

where

$$J(s_1, \dots, s_{j+1}, m_1, \dots, m_{k+1}) =$$

$$\frac{(-1)^{k+j}}{4\pi^2} \int_{L_A} \int_{L_B} \frac{f(z, w) dz dw}{(\lambda_{s_1}(A) - z) \cdots (\lambda_{s_{j+1}}(A) - z)(\lambda_{m_1}(B) - w) \cdots (\lambda_{m_{k+1}}(B) - w)}.$$

Lemma 2.5.1 from [7] gives us the estimate

$$\|C_{jk}\| \leq \max_{1 \leq s_1 < \dots < s_{j+1} < m_1 < \dots < m_{k+1} < n} |J(s_1, \dots, s_{j+1}, m_1, \dots, m_{k+1})| \| |V_A|^j |V_B|^k \|.$$

Due to Lemma 2.1,

$$|J(s_1, \dots, s_{j+1}, m_1, \dots, m_{k+1})| \leq \sup_{z \in co(A), w \in co(B)} \frac{|f^{(j,k)}(z, w)|}{j!k!}.$$

This inequality and (2.5) imply the required result.  $\square$

*Proof of Theorem 1.1:* Theorem 2.5.1 from [7] implies

$$(2.6) \quad \|V^k\| \leq \eta_k k! N^k(V)$$

for any nilpotent matrix  $V \in M(\mathbb{C}^n)$ . Take into account that  $N(|V_A|) = N(V_A)$ . Moreover, thanks to Lemma 2.3.2 from [7],  $N(V_A) = g(A)$ . Thus

$$(2.7) \quad \| |V_A|^k \| \leq k! \eta_k g^k(A) \quad (k = 1, \dots, n-1).$$

The similar inequality holds for  $V_B$ . Now the previous lemma yields the required result.  $\square$

**3. Functions of matrices on tensor products.** Let  $E_1 = \mathbb{C}^{n_1}$ ,  $E_2 = \mathbb{C}^{n_2}$ , be the Euclidean spaces of the dimensions  $n_1$  and  $n_2$ , with the scalar products  $\langle \cdot, \cdot \rangle_1$  and  $\langle \cdot, \cdot \rangle_2$ , respectively, and the norms  $\|\cdot\|_l = \sqrt{\langle \cdot, \cdot \rangle_l}$  ( $l = 1, 2$ ). Let  $H = E_1 \otimes E_2$  be the tensor product of  $E_1$  and  $E_2$  with the scalar product defined by

$$\langle y \otimes h, y_1 \otimes h_1 \rangle_H \equiv \langle y, y_1 \rangle_1 \langle h, h_1 \rangle_2 \quad (y, y_1 \in E_1; h, h_1 \in E_2)$$

and the cross norm  $\|\cdot\|_H = \sqrt{\langle \cdot, \cdot \rangle_H}$ , cf. [9]. In addition,  $I = I_H$  and  $I_l$  mean the unit operators in  $H$  and  $E_l$ , respectively. So  $H = \mathbb{C}^n$  with  $n = n_1 n_2$ .

Recall that  $M(E)$  is the set of all linear operators in a space  $E$ . In this section it is assumed that  $A \in M(E_1)$  and  $B \in M(E_2)$ .

Again let  $\Omega_A$  and  $\Omega_B$  be open simple connected sets containing  $\sigma(A)$  and  $\sigma(B)$ , respectively. Let  $f$  be a scalar function analytic on  $\Omega_A \times \Omega_B$ . Introduce the operator valued function

$$(3.1) \quad f(A, B) := -\frac{1}{4\pi^2} \int_{L_B} \int_{L_A} f(z, w) R_z(A) \otimes R_w(B) dw dz,$$

where  $L_A \subset \Omega_A, L_B \subset \Omega_B$  are closed contour surrounding  $\sigma(A)$  and  $\sigma(B)$ , respectively. If the series

$$f(z, w) = \sum_{k=0}^{\infty} \sum_{j=0}^{\infty} c_{jk} z^j w^k$$

converges for  $|z| \leq r_s(A), |w| \leq r_s(B)$ , where  $r_s(A)$  is the spectral radius of  $A$ , then (3.1) holds. Besides,

$$f(z, w) = \sum_{k=0}^{\infty} \sum_{j=0}^{\infty} c_{jk} A^j \otimes B^k.$$

**THEOREM 3.1.** *Let  $A \in M(E_1)$  and  $B \in M(E_2)$  and  $f(z, w)$  be regular on a neighborhood of  $co(A) \times co(B)$ . Then*

$$\|f(A, B)\|_H \leq \sum_{j=0}^{n_1-1} \sum_{k=0}^{n_2-1} \eta_j \eta_k g^j(A) g^k(B) \sup_{z \in co(A), w \in co(B)} |f^{(j,k)}(z, w)|.$$

*Proof.* Put  $\tilde{A} = A \otimes I_2, \tilde{B} = I_1 \otimes B_2$ . Now Lemma 2.2 implies

$$\|f(A, B)\|_H \leq \sum_{j,k=0}^{j+k \leq n-1} \sup_{z \in co(A), w \in co(B)} |f^{(j,k)}(z, w)| \frac{\| |V_{\tilde{A}}|^j |V_{\tilde{B}}|^k \|_H}{j!k!}$$

where  $V_{\tilde{A}}, V_{\tilde{B}}$  are the nilpotent parts of  $\tilde{A}$  and  $\tilde{B}$ , respectively. But

$$V_{\tilde{A}}^{n_1} = V_A^{n_1} \otimes I_2 = 0.$$

Similarly,  $V_{\tilde{B}}^{n_2} = 0$ . Thus,

$$\|f(A, B)\|_H \leq \sum_{k=0}^{n_2} \sum_{j=0}^{n_1} \sup_{z \in co(A), w \in co(B)} |f^{(j,k)}(z, w)| \frac{\| |V_A|^j \|_1 \| |V_B|^k \|_2}{j!k!}$$

Now the required result follows from (2.7).  $\square$

Taking into account (1.3), we get

**COROLLARY 3.2.** *Under the hypothesis of Theorem 3.1, the estimate*

$$\|f(A, B)\|_H \leq \sum_{k=0}^{n_2-1} \sum_{j=0}^{n_1-1} \frac{g^j(A) g^k(B)}{(j!k!)^{3/2}} \sup_{z \in co(A), w \in co(B)} |f^{(j,k)}(z, w)|$$

is true.



**Acknowledgment.** I am very grateful to the referee for very helpful remarks.

#### REFERENCES

- [1] R. Arens and A.P. Calderon. Analytic functions of several Banach algebra elements. *Ann. Math.*, 62:204–216, 1955.
- [2] Yu L. Daleckii and M.G. Krein. *Stability of Solutions of Differential Equations in Banach Space*. Amer. Math. Soc., Providence, R.I., 1974.
- [3] M.I. Gel'fand and G.E. Shilov. *Some Questions of Theory of Differential Equations*. Nauka, Moscow (in Russian), 1958.
- [4] A.O. Gel'fond. *Calculations of Finite Differences*. Nauka, Moscow (in Russian), 1967.
- [5] M.I. Gil'. Estimates for norm of matrix-valued functions, *Linear and Multilinear Algebra*, 35:65–73, 1993.
- [6] M.I. Gil'. *Stability of Finite and Infinite Dimensional Systems*. Kluwer Academic Publishers, Boston-Dordrecht-London, 1998.
- [7] M.I. Gil'. *Operator Functions and Localization of Spectra*. Lectures Notes In Mathematics vol. 1830, Springer-Verlag, Berlin, 2003.
- [8] R.A. Horn and C.R. Johnson. *Topics in Matrix Analysis*. Cambridge University Press, Cambridge, 1991.
- [9] M. Marcus. *Finite Dimensional Multilinear Algebra*, Vol. I, Marcel Dekker, New York, 1973.
- [10] V. Müller. *Spectral Theory of Linear Operators*. Birkhäuser Verlag, Basel, 2003.
- [11] A.M. Ostrowski. *Solution of Equations in Euclidean and Banach spaces*. Academic Press, New York - London, 1973.
- [12] J.L. Taylor. Analytic functional calculus for several commuting operators. *Acta Math.*, 125:1–38, 1970.