Chapter 2

Some Results of the Matrix Theory

This chapter is devoted to norm estimates for matrix-valued functions, in particular, for resolvents. These estimates will be applied in the rest of the book chapters.

In Section 2.1 we introduce the notation used in this chapter. In Section 2.2 we recall the well-known representations of matrix-valued functions. In Sections 2.3 and 2.4 we collect inequalities for the resolvent and present some results on spectrum perturbations of matrices. Sections 2.5 and 2.6 are devoted to matrix functions regular on simply-connected domains containing the spectrum. Section 2.7 deals with functions of matrices having geometrically simple eigenvalues; i.e., so-called diagonalizable matrices. In the rest of the chapter we consider particular cases of functions and matrices.

2.1 Notations

Everywhere in this chapter $\|x\|$ is the Euclidean norm of $x \in \mathbb{C}^n$: $\|x\| = \sqrt{(x,x)}$ with a scalar product $(.,.) = (.,.)_{C^n}$, I is the unit matrix.

For a linear operator A in $\mathbb{C}^n$ (matrix), $\lambda_k = \lambda_k(A)$ $(k = 1, \dots, n)$ are the eigenvalues of A enumerated in an arbitrary order with their multiplicities, $\sigma(A)$ denotes the spectrum of A, A^* is the adjoint to A, and A^{-1} is the inverse to A; $R_\lambda(A) = (A - \lambda I)^{-1}$ $(\lambda \in \mathbb{C}, \lambda \notin \sigma(A))$ is the resolvent, $r_s(A)$ is the spectral radius, $\|A\| = \sup_{x \in \mathbb{C}^n} \|Ax\|/\|x\|$ is the (operator) spectral norm, $N_2(A)$ is the Hilbert–Schmidt (Frobenius) norm of A: $N_2^2(A) = \operatorname{Trace} AA^*$, $A_I = (A - A^*)/2i$ is the imaginary component, $A_R = (A + A^*)/2$ is the real component,

$$\rho(A, \lambda) = \min_{k=1,\dots,n} |\lambda - \lambda_k(A)|$$

is the distance between $\sigma(A)$ and a point $\lambda \in \mathbb{C}$; $\rho(A, C)$ is the Hausdorff distance between a contour C and $\sigma(A)$. $\operatorname{co}(A)$ denotes the closed convex hull of $\sigma(A)$,

M.I. Gil', *Stability of Vector Differential Delay Equations*, Frontiers in Mathematics,
DOI 10.1007/978-3-0348-0577-3_2, © Springer Basel 2013

$\alpha(A) = \max_k \operatorname{Re} \lambda_k(A)$, $\beta(A) = \min_k \operatorname{Re} \lambda_k(A)$; $r_l(A)$ is the lower spectral radius:

$$r_l(A) = \min_{k=1,\dots,n} |\lambda_k(A)|.$$

In addition, $\mathbb{C}^{n\times n}$ is the set of complex $n \times n$-matrices.

The following quantity plays an essential role in the sequel:

$$g(A) = \left(N_2^2(A) - \sum_{k=1}^{n} |\lambda_k(A)|^2\right)^{1/2}.$$

It is not hard to check that

$$g^2(A) \le N_2^2(A) - |\operatorname{Trace} A^2|.$$

In Section 2.2 of the book [31] it is proved that

$$g^2(A) \le 2N_2^2(A_I) \tag{1.1}$$

and

$$g(e^{i\tau}A + zI) = g(A) \tag{1.2}$$

for all $\tau \in \mathbb{R}$ and $z \in \mathbb{C}$.

2.2 Representations of matrix functions

2.2.1 Classical representations

In this subsection we recall some classical representations of functions of matrices. For details see [74, Chapter 6] and [9].

Let $A \in \mathbb{C}^{n\times n}$ and $M \supset \sigma(A)$ be an open simply-connected set whose boundary C consists of a finite number of rectifiable Jordan curves, oriented in the positive sense customary in the theory of complex variables. Suppose that $M \cup C$ is contained in the domain of analyticity of a scalar-valued function f. Then $f(A)$ can be defined by the generalized integral formula of Cauchy

$$f(A) = -\frac{1}{2\pi i}\int_C f(\lambda) R_\lambda(A)\, d\lambda. \tag{2.1}$$

If an analytic function $f(\lambda)$ is represented by the Taylor series

$$f(\lambda) = \sum_{k=0}^{\infty} c_k \lambda^k \quad \left(|\lambda| < \frac{1}{\overline{\lim}_{k\to\infty} \sqrt[k]{|c_k|}}\right),$$

then one can define $f(A)$ as

$$f(A) = \sum_{k=0}^{\infty} c_k A^k$$

provided the spectral radius $r_s(A)$ of A satisfies the inequality

$$r_s(A) \lim_{k\to\infty} \sqrt[k]{|c_k|} < 1.$$

In particular, for any matrix A,

$$e^A = \sum_{k=0}^{\infty} \frac{A^k}{k!}.$$

Consider the $n \times n$-Jordan block:

$$J_n(\lambda_0) = \begin{pmatrix} \lambda_0 & 1 & 0 & \dots & 0 \\ 0 & \lambda_0 & 1 & \dots & 0 \\ . & . & . & \dots & . \\ . & . & . & \dots & . \\ . & . & . & \dots & . \\ 0 & 0 & \dots & \lambda_0 & 1 \\ 0 & 0 & \dots & 0 & \lambda_0 \end{pmatrix},$$

then

$$f(J_n(\lambda_0)) = \begin{pmatrix} f(\lambda_0) & \frac{f'(\lambda_0)}{1!} & \dots & \frac{f^{(n-1)}(\lambda_0)}{(n-1)!} \\ 0 & f(\lambda_0) & \dots & \\ . & . & \dots & . \\ . & . & \dots & . \\ . & . & \dots & . \\ 0 & \dots & f(\lambda_0) & \frac{f'(\lambda_0)}{1!} \\ 0 & \dots & 0 & f(\lambda_0) \end{pmatrix}.$$

Thus, if A has the Jordan block-diagonal form

$$A = \operatorname{diag}(J_{m_1}(\lambda_1), J_{m_2}(\lambda_2), \dots, J_{m_{n_0}}(\lambda_{n_0})),$$

where λ_k, $k = 1, \dots, n_0$ are the eigenvalues whose geometric multiplicities are m_k, then

$$f(A) = \operatorname{diag}(f(J_{m_1}(\lambda_1)), f(J_{m_2}(\lambda_2)), \dots, f(J_{m_{n_0}}(\lambda_{n_0}))). \tag{2.2}$$

Note that in (2.2) we do not require that f is regular on a neighborhood of an open set containing $\sigma(A)$; one can take an arbitrary function which has at each λ_k derivatives up to $m_k - 1$-order.

In particular, if an $n \times n$-matrix A is diagonalizable, that is its eigenvalues have the geometric multiplicities $m_k \equiv 1$, then

$$f(A) = \sum_{k=1}^{n} f(\lambda_k) Q_k, \tag{2.3}$$

where Q_k are the eigenprojections. In the case (2.3) it is required only that f is defined on the spectrum.

Now let

$$\sigma(A) = \cup_{k=1}^{m} \sigma_k(A) \ \ (m \leq n)$$

and $\sigma_k(A) \subset M_k$ $(k = 1, \ldots, m)$, where M_k are open disjoint simply-connected sets: $M_k \cap M_j = \emptyset$ $(j \neq k)$. Let f_k be regular on M_k. Introduce on $M = \cup_{k=1}^{m} M_k$ the piece-wise analytic function by $f(z) = f_j(z)$ $(z \in M_j)$. Then

$$f(A) = -\frac{1}{2\pi i} \sum_{j=1}^{m} \int_{C_j} f(\lambda) R_\lambda(A) d\lambda, \tag{2.4}$$

where $C_j \subset M_j$ are closed smooth contours surrounding $\sigma(A_j)$ and the integration is performed in the positive direction. For more details about representation (2.4) see [110, p. 49].

For instance, let M_1 and M_2 be two disjoint disks, and

$$f(z) = \begin{cases} \sin z & \text{if } z \in M_1, \\ \cos z & \text{if } z \in M_2. \end{cases}$$

Then (2.4) holds with $m = 2$.

2.2.2 Multiplicative representations of the resolvent

In this subsection we present multiplicative representations of the resolvent which lead to new representations of matrix functions.

Let $A \in \mathbb{C}^{n \times n}$ and λ_k be its eigenvalues with the multiplicities taken into account.

As it is well known, there is an orthogonal normal basis (the Schur basis) $\{e_k\}$ in which A is represented by a triangular matrix. Moreover there is the (maximal) chain P_k $(k = 1, \ldots, n)$ of the invariant orthogonal projections of A. That is, $AP_k = P_k A P_k$ $(k = 1, \ldots, n)$ and

$$0 = P_0 \mathbb{C}^n \subset P_1 \mathbb{C}^n \subset \cdots \subset P_n \mathbb{C}^n = \mathbb{C}^n.$$

So $\dim(P_k - P_{k-1})\mathbb{C}^n = 1$. Besides,

$$A = D + V \ \ (\sigma(A) = \sigma(D)), \tag{2.5}$$

where

$$D = \sum_{k=1}^{n} \lambda_k \Delta P_k \ \ (\Delta P_k = P_k - P_{k-1}) \tag{2.6}$$

is the diagonal part of A and V is the nilpotent part of A. That is, V is a nilpotent matrix, such that

$$VP_k = P_{k-1} A P_k \ (k = 2, \ldots, n). \tag{2.7}$$

For more details see for instance, [18]. The representation (2.5) will be called the triangular (Schur) representation.

Furthermore, for $X_1, X_2, \ldots, X_j \in \mathbb{C}^{n\times n}$ let

$$\prod_{1\le k\le j}^{\rightarrow} X_k \equiv X_1 X_2 \ldots X_n.$$

That is, the arrow over the symbol of the product means that the indexes of the co-factors increase from left to right.

Theorem 2.2.1. *Let D and V be the diagonal and nilpotent parts of an $A \in \mathbb{C}^{n\times n}$, respectively. Then*

$$R_\lambda(A) = R_\lambda(D) \prod_{2\le k\le n}^{\rightarrow} \left[I + \frac{V\Delta P_k}{\lambda - \lambda_k}\right] \quad (\lambda \notin \sigma(A)),$$

where $P_k,\ k = 1, \ldots, n$, is the maximal chain of the invariant projections of A.

For the proof see [31, Theorem 2.9.1]. Since

$$R_\lambda(D) = \sum_{j=1}^{n} \frac{\Delta P_k}{\lambda_j - \lambda},$$

from the previous theorem we have the following result.

Corollary 2.2.2. *The equality*

$$R_\lambda(A) = \sum_{j=1}^{n} \frac{\Delta P_j}{\lambda_j - \lambda} \prod_{j+1\le k\le n+1}^{\rightarrow} \left[I + \frac{V\Delta P_k}{\lambda - \lambda_k}\right] \quad (\lambda \notin \sigma(A))$$

is true with $V\Delta P_{n+1} = 0$.

Now (2.1) implies the representation

$$f(A) = -\frac{1}{2\pi i}\int_C f(\lambda) \sum_{j=1}^{n} \frac{\Delta P_j}{\lambda_j - \lambda} \prod_{j+1\le k\le n+1}^{\rightarrow} \left[I + \frac{V\Delta P_k}{\lambda - \lambda_k}\right] d\lambda.$$

Here one can apply the residue theorem.

Furthermore, the following result is proved in [31, Theorem 2.9.1].

Theorem 2.2.3. *For any $A \in \mathbb{C}^{n\times n}$ we have*

$$\lambda R_\lambda = -\prod_{1\le k\le n}^{\rightarrow} \left(I + \frac{A\Delta P_k}{\lambda - \lambda_k}\right) \quad (\lambda \notin \sigma(A) \cup 0), \tag{2.8}$$

where $P_k,\ k = 1, \ldots, n$ is the maximal chain of the invariant projections of A.

Let A be a normal matrix. Then

$$A = \sum_{k=1}^{n} \lambda_k \Delta P_k.$$

Hence, $A\Delta P_k = \lambda_k \Delta P_k$. Since $\Delta P_k \Delta P_j = 0$ for $j \neq k$, the previous theorem gives us the equality

$$-\lambda R_\lambda(A) = I + \sum_{k=1}^{n} \frac{\lambda_k \Delta P_k}{\lambda - \lambda_k} = \sum_{k=1}^{n} \left(\Delta P_k + \frac{\lambda_k \Delta P_k}{\lambda - \lambda_k} \right)$$

or

$$R_\lambda(A) = \sum_{k=1}^{n} \frac{\Delta P_k}{\lambda_k - \lambda}.$$

So from (2.8) we have obtained the well-known spectral representation for the resolvent of a normal matrix. Thus, the previous theorem generalizes the spectral representation for the resolvent of a normal matrix. Now we can use (2.1) and (2.8) to get the representation for $f(A)$.

2.3 Norm estimates for resolvents

For a natural $n \geq 2$, introduce the numbers

$$\gamma_{n,k} = \sqrt{\frac{(n-1)(n-2)\dots(n-k)}{(n-1)^k k!}} \quad (k = 1, 2, \dots, n-1),\ \gamma_{n,0} = 1.$$

Evidently, for all $n > 2$,

$$\gamma_{n,k}^2 \leq \frac{1}{k!} \quad (k = 1, 2, \dots, n-1). \tag{3.1}$$

Theorem 2.3.1. *Let A be a linear operator in $\mathbb{C}^n$. Then its resolvent satisfies the inequality*

$$\|R_\lambda(A)\| \leq \sum_{k=0}^{n-1} \frac{g^k(A)\gamma_{n,k}}{\rho^{k+1}(A,\lambda)} \quad (\lambda \notin \sigma(A)),$$

where $\rho(A,\lambda) = \min_{k=1,\dots,n} |\lambda - \lambda_k(A)|$.

To prove this theorem again use the Schur triangular representation (2.5). Recall $g(A)$ is defined in Section 2.1. As it is shown in [31, Section 2.1]), the relation $g(U^{-1}AU) = g(A)$ is true, if U is a unitary matrix. Hence it follows that $g(A) = N_2(V)$. The proof of the previous theorem is based on the following lemma.

Lemma 2.3.2. *The inequality*

$$\|R_\lambda(A) - R_\lambda(D)\| \le \sum_{k=1}^{n-1} \frac{\gamma_{n,k} g^k(A)}{\rho^{k+1}(A,\lambda)} \quad (\lambda \notin \sigma(A))$$

is true.

Proof. By (2.5) we have

$$R_\lambda(A) - R_\lambda(D) = -R_\lambda(D) V R_\lambda(A) = -R_\lambda(D) V (D + V - I\lambda)^{-1}.$$

Thus

$$R_\lambda(A) - R_\lambda(D) = -R_\lambda(D) V (I + R_\lambda(D) V)^{-1} R_\lambda(D). \tag{3.2}$$

Clearly, $R_\lambda(D)V$ is a nilpotent matrix. Hence,

$$R_\lambda(A) - R_\lambda(D) = -R_\lambda(D) V \sum_{k=0}^{n-1} (-R_\lambda(D)V)^k R_\lambda(D) = \sum_{k=1}^{n-1} (-R_\lambda(D)V)^k R_\lambda(D). \tag{3.3}$$

Thanks to Theorem 2.5.1 [31], for any $n \times n$ nilpotent matrix V_0,

$$\|V_0^k\| \le \gamma_{n,k} N_2^k(V_0). \tag{3.4}$$

In addition, $\|R_\lambda(D)\| = \rho^{-1}(D,\lambda) = \rho^{-1}(A,\lambda)$. So

$$\|(-R_\lambda(D)V)^k\| \le \gamma_{n,k} N_2^k(R_\lambda(D)V) \le \frac{\gamma_{n,k} N_2^k(V)}{\rho^k(A,\lambda)}.$$

Now the required result follows from (3.3). □

The assertion of Theorem 2.3.1 directly follows from the previous lemma.

Note that the just proved Lemma 2.3.2 is a slight improvement of Theorem 2.1.1 from [31].

Theorem 2.3.1 is sharp: if A is a normal matrix, then $g(A) = 0$ and Theorem 2.3.1 gives us the equality $\|R_\lambda(A)\| = 1/\rho(A,\lambda)$. Taking into account (3.1), we get

Corollary 2.3.3. *Let $A \in \mathbb{C}^{n\times n}$. Then*

$$\|R_\lambda(A)\| \le \sum_{k=0}^{n-1} \frac{g^k(A)}{\sqrt{k!}\,\rho^{k+1}(A,\lambda)}$$

for any regular λ of A.

We will need the following result.

Theorem 2.3.4. *Let $A \in \mathbb{C}^{n\times n}$. Then*

$$\|R_\lambda(A)\ \det(\lambda I - A)\| \le \left[\frac{N_2^2(A) - 2\,\mathrm{Re}(\overline{\lambda}\ \mathrm{Trace}(A))\ + n|\lambda|^2}{n-1}\right]^{(n-1)/2} \quad (\lambda \notin \sigma(A)).$$

In particular, let V be a nilpotent matrix. Then

$$\|(I\lambda - V)^{-1}\| \le \frac{1}{|\lambda|}\left[1 + \frac{1}{n-1}\left(1 + \frac{N_2^2(V)}{|\lambda|^2}\right)\right]^{(n-1)/2} \quad (\lambda \ne 0).$$

The proof of this theorem can be found in [31, Section 2.11].

We also point to the following result.

Theorem 2.3.5. *Let $A \in \mathbb{C}^{n\times n}$. Then*

$$\|R_\lambda(A)\| \le \frac{1}{\rho(A,\lambda)}\left[1 + \frac{1}{n-1}\left(1 + \frac{g^2(A)}{\rho^2(A,\lambda)}\right)\right]^{(n-1)/2}$$

for any regular λ of A.

For the proof see [31, Theorem 2.14.1].

2.4 Spectrum perturbations

Let A and B be $n \times n$-matrices having eigenvalues

$$\lambda_1(A), \dots, \lambda_n(A) \quad \text{and} \quad \lambda_1(B), \dots, \lambda_n(B),$$

respectively, and $q = \|A - B\|$.

The spectral variation of B with respect to A is

$$sv_A(B) := \max_i\ \min_j |\lambda_i(B) - \lambda_j(A)|,$$

cf. [103]. The following simple lemma is proved in [31, Section 4.1].

Lemma 2.4.1. *Assume that $\|R_\lambda(A)\| \le \phi(\rho^{-1}(A,\lambda))$ for all regular λ of A, where $\phi(x)$ is a monotonically increasing non-negative continuous function of a non-negative variable x, such that $\phi(0) = 0$ and $\phi(\infty) = \infty$. Then the inequality $sv_A(B) \le z(\phi, q)$ is true, where $z(\phi, q)$ is the unique positive root of the equation $q\phi(1/z) = 1$.*

This lemma and Corollary 2.3.2 yield our next result.

Theorem 2.4.2. *Let A and B be $n \times n$-matrices. Then $sv_A(B) \le z(q, A)$, where $z(q, A)$ is the unique non-negative root of the algebraic equation*

$$y^n = q \sum_{j=0}^{n-1} \frac{y^{n-j-1} g^j(A)}{\sqrt{j!}}. \tag{4.1}$$

Let us consider the algebraic equation

$$z^n = P(z) \quad (n > 1), \quad \text{where} \quad P(z) = \sum_{j=0}^{n-1} c_j z^{n-j-1} \tag{4.2}$$

with non-negative coefficients c_j $(j = 0, \dots, n-1)$.

Lemma 2.4.3. *The extreme right-hand root z_0 of equation (4.2) is non-negative and the following estimates are valid:*

$$z_0 \le \sqrt[n]{P(1)} \ \text{ if } \ P(1) \le 1, \tag{4.3}$$

and

$$1 \le z_0 \le P(1) \ \text{ if } \ P(1) \ge 1. \tag{4.4}$$

Proof. Since all the coefficients of $P(z)$ are non-negative, it does not decrease as $z > 0$ increases. From this it follows that if $P(1) \le 1$, then $z_0 \le 1$. So $z_0^n \le P(1)$, as claimed.

Now let $P(1) \ge 1$, then due to (4.2) $z_0 \ge 1$ because $P(z)$ does not decrease. It is clear that

$$P(z_0) \le z_0^{n-1} P(1)$$

in this case. Substituting this inequality into (4.2), we get (4.4). □

Substituting $z = ax$ with a positive constant a into (4.2), we obtain

$$x^n = \sum_{j=0}^{n-1} \frac{c_j}{a^{j+1}} x^{n-j-1}. \tag{4.5}$$

Let

$$a = 2 \max_{j=0,\dots,n-1} \sqrt[j+1]{c_j}.$$

Then

$$\sum_{j=0}^{n-1} \frac{c_j}{a^{j+1}} \le \sum_{j=0}^{n-1} 2^{-j-1} = 1 - 2^{-n} < 1.$$

Let x_0 be the extreme right-hand root of equation (4.5), then by (4.3) we have $x_0 \le 1$. Since $z_0 = ax_0$, we have derived the following result.

Corollary 2.4.4. *The extreme right-hand root z_0 of equation* (4.2) *is non-negative. Moreover,*

$$z_0 \le 2 \max_{j=0,\dots,n-1} \sqrt[j+1]{c_j}.$$

Now put $y = xg(A)$ into (6.1). Then we obtain the equation

$$x^n = \frac{q}{g(A)} \sum_{j=0}^{n-1} \frac{x^{n-j-1}}{\sqrt{j!}}.$$

Put

$$w_n = \sum_{j=0}^{n-1} \frac{1}{\sqrt{j!}}.$$

Applying Lemma 2.4.3 we get the estimate $z(q, A) \le \delta(q)$, where

$$\delta(q) := \begin{cases} q w_n & \text{if } q w_n \ge g(A), \\ g^{1-1/n}(A)[q w_n]^{1/n} & \text{if } q w_n \le g(A). \end{cases}$$

Now Theorem 2.4.2 ensures the following result.

Corollary 2.4.5. *One has $sv_A(B) \le \delta(q)$.*

Furthermore let $\tilde{D}, V_+$ and V_- be the diagonal, upper nilpotent part and lower nilpotent part of matrix A, respectively. Using the notation $A_+ = \tilde{D} + V_+$, we arrive at the relations

$$\sigma(A_+) = \sigma(\tilde{D}), g(A_+) = N_2(V_+) \quad \text{and} \quad \|A - A_+\| = \|V_-\|.$$

Taking

$$\delta_A := \begin{cases} \|V_-\| w_n & \text{if } \|V_-\| w_n \ge N_2(V_+), \\ N_2^{1-1/n}(V_+)[\|V_-\| w_n]^{1/n} & \text{if } \|V_-\| w_n \le N_2(V_+), \end{cases}$$

due to the previous corollary we obtain

Corollary 2.4.6. *Let $A = (a_{jk})_{j,k=1}^n$ be an $n \times n$-matrix. Then for any eigenvalue μ of A, there is a $k = 1, \dots, n$, such that*

$$|\mu - a_{kk}| \le \delta_A,$$

and therefore the (*upper*) *spectral radius satisfies the inequality*

$$r_s(A) \le \max_{k=1,\dots,n} |a_{kk}| + \delta_A,$$

and the lower spectral radius satisfies the inequality

$$r_l(A) \ge \min_{k=1,\dots,n} |a_{kk}| - \delta_A,$$

provided $|a_{kk}| > \delta_A$ $(k = 1, \dots, n)$.

Clearly, one can exchange the places V_+ and V_-.

Let us recall the celebrated Gerschgorin theorem. To this end write

$$R_j = \sum_{k=1, k\neq j}^{n} |a_{jk}|.$$

Let $\Omega(b, r)$ be the closed disc centered at $b \in \mathbb{C}$ with a radius r.

Theorem 2.4.7 (Gerschgorin)**.** *Every eigenvalue of A lies within at least one of the discs $\Omega(a_{jj}, R_j)$.*

Proof. Let λ be an eigenvalue of A and let $x = (x_j)$ be the corresponding eigenvector. Let i be chosen so that $|x_i| = \max_j |x_j|$. Then $|x_i| > 0$, otherwise $x = 0$. Since x is an eigenvector, $Ax = \lambda x$ or equivalent

$$\sum_{k=1}^{n} a_{ik} = \lambda x_i$$

so, splitting the sum, we get

$$\sum_{k=1, k\neq i}^{n} a_{ik} x_k = \lambda x_i - a_{ii} x_i.$$

We may then divide both sides by x_i (choosing i as we explained we can be sure that $x_i \neq 0$) and take the absolute value to obtain

$$|\lambda - a_{ii}| \leq \sum_{k=1, k\neq i}^{n} |a_{jk}| \frac{|x_k|}{|x_i|} \leq R_i,$$

where the last inequality is valid because

$$\frac{|x_k|}{|x_l|} \leq 1$$

as claimed. $\square$

Note that for a diagonal matrix the Gerschgorin discs $\Omega(a_{jj}, R_j)$ coincide with the spectrum. Conversely, if the Gerschgorin discs coincide with the spectrum, the matrix is diagonal.

The next lemma follows from the Gerschgorin theorem and gives us a simple bound for the spectral radius.

Lemma 2.4.8. *The spectral radius $r_s(A)$ of a matrix $A = (a_{jk})_{j,k=1}^{n}$ satisfies the inequality*

$$r_s(A) \leq \max_j \sum_{k=1}^{n} |a_{jk}|.$$

About this and other estimates for the spectral radius see [80, Section 16].

2.5 Norm estimates for matrix functions

2.5.1 Estimates via the resolvent

The following result directly follows from (2.1).

Lemma 2.5.1. *Let let $f(\lambda)$ be a scalar-valued function which is regular on a neighborhood M of an open simply-connected set containing the spectrum of $A \in \mathbb{C}^{n\times n}$, and $C \subset M$ be a closed smooth contour surrounding $\sigma(A)$. Then*

$$\|f(A)\| \le \frac{1}{2\pi}\int_C |f\ (z)|\|R_z(A)\|dz \le m_C(A) l_C \sup_{z\in C}|f(z)|,$$

where

$$m_C(A) := \sup_{z\in C}\|R_z(A)\|,\ l_C := \frac{1}{2\pi}\int_C |dz|.$$

Now we can directly apply the estimates for the resolvent from Section 2.3. In particular, by Corollary 2.3.3 we have

$$\|R_z(A)\| \le p(A, 1/\rho(A,z)), \tag{5.1}$$

where

$$p(A,x) = \sum_{k=0}^{n-1}\frac{x^{k+1}g^k(A)}{\sqrt{k!}}\quad (x>0). \tag{5.2}$$

We thus get $m_C(A) \le p(A, 1/\rho(A,C))$, where $\rho(A,C)$ is the distance between C and $\sigma(A)$, and therefore,

$$\|f(A)\| \le l_C p(A, 1/\rho(A,C))\sup_{z\in C}|f\ (z)|. \tag{5.3}$$

2.5.2 Functions regular on the convex hull of the spectrum

Theorem 2.5.2. *Let A be an $n\times n$-matrix and f be a function holomorphic on a neighborhood of the convex hull* $\mathrm{co}(A)$ *of $\sigma(A)$. Then*

$$\|f(A)\| \le \sup_{\lambda\in\sigma(A)}|f(\lambda)| + \sum_{k=1}^{n-1}\sup_{\lambda\in \mathrm{co}(A)}|f^{(k)}(\lambda)|\frac{\gamma_{n,k}g^k(A)}{k!}.$$

This theorem is proved in the next subsection. Taking into account (3.1) we get our next result.

Corollary 2.5.3. *Under the hypothesis of Theorem* 2.5.2 *we have*

$$\|f(A)\| \le \sup_{\lambda\in\sigma(A)}|f(\lambda)| + \sum_{k=1}^{n-1}\sup_{\lambda\in \mathrm{co}(A)}|f^{(k)}(\lambda)|\frac{g^k(A)}{(k!)^{3/2}}.$$

Theorem 2.5.2 is sharp: if A is normal, then $g(A) = 0$ and

$$\|f(A)\| = \sup_{\lambda \in \sigma(A)} |f(\lambda)|.$$

For example,

$$\|\exp(At)\| \leq e^{\alpha(A)t} \sum_{k=0}^{n-1} g^k(A) t^k \frac{\gamma_{n,k}}{k!} \leq e^{\alpha(A)t} \sum_{k=0}^{n-1} \frac{g^k(A) t^k}{(k!)^{3/2}} \quad (t \geq 0)$$

where $\alpha(A) = \max_{k=1,\dots,n} \operatorname{Re} \lambda_k(A)$. In addition,

$$\|A^m\| \leq \sum_{k=0}^{n-1} \frac{\gamma_{n,k} m! g^k(A) r_s^{m-k}(A)}{(m-k)!k!} \leq \sum_{k=0}^{n-1} \frac{m! g^k(A) r_s^{m-k}(A)}{(m-k)!(k!)^{3/2}} \quad (m = 1, 2, \dots),$$

where $r_s(A)$ is the spectral radius. Recall that $1/(m-k)! = 0$ if $m < k$.

2.5.3 Proof of Theorem 2.5.2

Let $|V|_e$ be the operator whose entries in the orthonormal basis of the triangular representation (the Schur basis) $\{e_k\}$ are the absolute values of the entries of the nilpotent part V of A with respect to this basis. That is,

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

where $a_{jk} = (Ae_k, e_j)$. Put

$$I_{j_1 \dots j_{k+1}} = \frac{(-1)^{k+1}}{2\pi i} \int_C \frac{f(\lambda) d\lambda}{(\lambda_{j_1} - \lambda) \dots (\lambda_{j_{k+1}} - \lambda)}.$$

We need the following result.

Lemma 2.5.4. *Let A be an $n \times n$-matrix and f be a holomorphic function on a Jordan domain (that is on a closed simply connected set, whose boundary is a Jordan contour), containing $\sigma(A)$. Let D be the diagonal part of A. Then*

$$\|f(A) - f(D)\| \leq \sum_{k=1}^{n-1} J_k \| \, |V|_e^k \|,$$

where

$$J_k = \max\{|I_{j_1 \dots j_{k+1}}| : 1 \leq j_1 < \dots < j_{k+1} \leq n\}.$$

Proof. From (2.1) and (3.3) we deduce that

$$f(A) - f(D) = -\frac{1}{2\pi i}\int_C f(\lambda)(R_\lambda(A) - R_\lambda(D))d\lambda = \sum_{k=1}^{n-1} B_k, \tag{5.4}$$

where

$$B_k = (-1)^{k+1}\frac{1}{2\pi i}\int_C f(\lambda)(R_\lambda(D)V)^k R_\lambda(D)d\lambda.$$

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

$$R_\lambda(D) = \sum_{j=1}^{n}\frac{\Delta P_j}{\lambda_j(A) - \lambda},$$

where $\Delta P_k = (., e_k)e_k$. In addition, $\Delta P_j V \Delta P_k = 0$ for $j \geq k$. Consequently,

$$B_k = \sum_{j_1=1}^{j_2-1}\Delta P_{j_1} V \sum_{j_2=1}^{j_3-1}\Delta P_{j_2} V \cdots \sum_{j_k=1}^{j_{k+1}-1} V \sum_{j_{k+1}=1}^{n}\Delta P_{j_{k+1}} I_{j_1 j_2 \ldots j_{k+1}}.$$

Lemma 2.8.1 from [31] gives us the estimate

$$\begin{aligned}\|B_k\| &\leq J_k \|\sum_{j_1=1}^{j_2-1}\Delta P_{j_1}|V|_e \sum_{j_2=1}^{j_3-1}\Delta P_{j_2}|V|_e \ \cdots \sum_{j_k=1}^{j_{k+1}-1}|V|_e \sum_{j_{k+1}=1}^{n}\Delta P_{j_{k+1}}\| \\ &= J_k\|P_{n-k}|V|_e P_{n-k+1}|V|_e P_{n-k+2}\ldots P_{n-1}|V|_e\|.\end{aligned}$$

But

$$\begin{aligned}P_{n-k}|V|_e P_{n-k+1}|V|_e P_{n-k+2}\ldots P_{n-1}|V|_e &= |V|_e P_{n-k+1}|V|_e P_{n-k+2}\ldots P_{n-1}|V|_e \\ &= |V|_e^2 \ldots P_{n-1}|V|_e = |V|_e^k.\end{aligned}$$

Thus

$$\|B_k\| \leq J_k\| \ |V|_e^k \ \|.$$

This inequality and (5.4) imply the required inequality. □

Since $N_2(|V|_e) = N_2(V) = g(A)$, by (3.4) and the previous lemma we get the following result.

Lemma 2.5.5. *Under the hypothesis of Lemma* 5.4 *we have*

$$\|f(A) - f(D)\| \leq \sum_{k=1}^{n-1} J_k \gamma_{n,k} g^k(A).$$

Let f be holomorphic on a neighborhood of $\mathrm{co}(A)$. Thanks to Lemma 1.5.1 from [31],

$$J_k \le \frac{1}{k!} \sup_{\lambda\in \mathrm{co}(A)} |f^{(k)}(\lambda)|.$$

Now the previous lemma implies

Corollary 2.5.6. *Under the hypothesis of Theorem* 2.5.2 *we have the inequalities*

$$\|f(A) - f(D)\| \le \sum_{k=1}^{n-1} \sup_{\lambda\in \mathrm{co}(A)} |f^{(k)}(\lambda)|\gamma_{n,k} \frac{g^k(A)}{k!}.$$

The assertion of Theorem 2.5.2 directly follows from the previous corollary.

Note that the latter corollary is a slight improvement of Theorem 2.7.1 from [31].

Denote by $f[a_1, a_2, \dots, a_{k+1}]$ the kth divided difference of f at points a_1, a_2, $\dots$, a_{k+1}. By the Hadamard representation [20, formula (54)], we have

$$I_{j_1 \dots j_{k+1}} = f[\lambda_1, \dots, \lambda_{j_{k+1}}],$$

provided λ_j are distinct. Now Lemma 5.5 implies

Corollary 2.5.7. *Let all the eigenvalues of an $n\times n$-matrix A be algebraically simple, and f be a holomorphic function in a Jordan domain containing $\sigma(A)$. Then*

$$\|f(A) - f(D)\| \le \sum_{k=1}^{n-1} f_k \gamma_{n,k} g^k(A) \le \sum_{k=1}^{n-1} f_k \frac{g^k(A)}{\sqrt{k!}},$$

where

$$f_k = \max\{|f[\lambda_1(A), \dots, \lambda_{j_{k+1}}(A)]| : 1 < j_1 < \cdots < j_{k+1} \le n\}.$$

2.6 Absolute values of entries of matrix functions

Everywhere in the present section, $A = (a_{jk})_{j,k=1}^n$, $S = \mathrm{diag}[a_{11}, \dots, a_{nn}]$ and the off diagonal of A is $W = A - S$. That is, the entries v_{jk} of W are $v_{jk} = a_{jk}$ ($j \ne k$) and $v_{jj} = 0$ ($j,k = 1, 2, \dots$). Denote by $\mathrm{co}(S)$ the closed convex hull of the diagonal entries $a_{11}, \dots, a_{nn}$. We put $|A| = (|a_{jk}|)_{j,l=1}^n$, i.e., $|A|$ is the matrix whose entries are the absolute values of the entries A in the standard basis. We also write $T \ge 0$ if all the entries of a matrix T are non-negative. If T and B are two matrices, then we write $T \ge B$ if $T - B \ge 0$.

Thanks to Lemma 2.4.8 we obtain $r_s(|W|) \le \tau_W$, where

$$\tau_W := \max_j \sum_{k=1, k\ne j}^{n} |a_{jk}|.$$

Theorem 2.6.1. *Let $f(\lambda)$ be holomorphic on a neighborhood of a Jordan set, whose boundary C has the property*

$$|z - a_{jj}| > \sum_{k=1, k\neq j}^{n} |a_{jk}| \tag{6.1}$$

for all $z \in C$ and $j = 1, \dots, n$. Then, with the notation

$$\xi_k(A) := \sup_{z\in \text{co }(S)} \frac{|f^{(k)}(z)|}{k!} \quad (k = 1, 2, \dots),$$

the inequality

$$|f(A) - f(S)| \le \sum_{k=1}^{\infty} \xi_k(A)|W|^k$$

is valid, provided

$$r_s(|W|) \lim_{k\to\infty} \sqrt[k]{\xi_k(A)} < 1.$$

Proof. By the equality $A = S + W$ we get

$$R_\lambda(A) = (S+W-\lambda I)^{-1} = (I+R_\lambda(S)W)^{-1}R_\lambda(S) = \sum_{k=0}^{\infty}(R_\lambda(S)W)^k(-1)^k R_\lambda(S),$$

provided the spectral radius $r_0(\lambda)$ of $R_\lambda(S)W$ is less than one. The entries of this matrix are

$$\frac{a_{jk}}{a_{jj} - \lambda} \quad (\lambda \neq a_{jj}, \;\; j \neq k)$$

and the diagonal entries are zero. Thanks to Lemma 2.4.8 we have

$$r_0(\lambda) \le \max_j \sum_{k=1, k\neq j}^{n} \frac{|a_{jk}|}{|a_{jj} - \lambda|} < 1 \quad (\lambda \in C)$$

and the series

$$R_\lambda(A) - R_\lambda(S) = \sum_{k=1}^{\infty}(R_\lambda(S)W)^k(-1)^k R_\lambda(S)$$

converges. Thus

$$f(A) - f(S) = -\frac{1}{2\pi i}\int_C f(\lambda)(R_\lambda(A) - R_\lambda(S))d\lambda = \sum_{k=1}^{\infty} M_k, \tag{6.2}$$

where

$$M_k = (-1)^{k+1}\frac{1}{2\pi i}\int_C f(\lambda)(R_\lambda(S)W)^k R_\lambda(S)d\lambda.$$

Since S is a diagonal matrix with respect to the standard basis $\{d_k\}$, we can write

$$R_\lambda(S) = \sum_{j=1}^{n} \frac{\hat{Q}_j}{b_j - \lambda} \ \ (b_j = a_{jj}),$$

where $\hat{Q}_k = (., d_k)d_k$. We thus have

$$M_k = \sum_{j_1=1}^{n} \hat{Q}_{j_1} W \sum_{j_2=1}^{n} \hat{Q}_{j_2} W \dots W \sum_{j_{k+1}=1}^{n} \hat{Q}_{j_{k+1}} J_{j_1 j_2 \dots j_{k+1}}. \tag{6.3}$$

Here

$$J_{j_1 \dots j_{k+1}} = \frac{(-1)^{k+1}}{2\pi i} \int_C \frac{f(\lambda) d\lambda}{(b_{j_1} - \lambda) \dots (b_{j_{k+1}} - \lambda)}.$$

Lemma 1.5.1 from [31] gives us the inequalities

$$|J_{j_1 \dots j_{k+1}}| \le \xi_k(A) \ \ (j_1, j_2, \dots, j_{k+1} = 1, \dots, n).$$

Hence, by (6.3)

$$|M_k| \le \xi_k(A) \sum_{j_1=1}^{n} \hat{Q}_{j_1} |W| \sum_{j_2=1}^{n} \hat{Q}_{j_2} |W| \dots |W| \sum_{j_{k+1}=1}^{n} \hat{Q}_{j_{k+1}}.$$

But

$$\sum_{j_1=1}^{n} \hat{Q}_{j_1} |W| \sum_{j_2=1}^{n} \hat{Q}_{j_2} |W| \dots |W| \sum_{j_{k+1}=1}^{n} \hat{Q}_{j_{k+1}} = |W|^k.$$

Thus $|M_k| \le \xi_k(A)|W|^k$. Now (6.2) implies the required result. □

Additional estimates for the entries of matrix functions can be found in [43, 38]. Under the hypothesis of the previous theorem with the notation

$$\xi_0(A) := \max_k |f(a_{kk})|,$$

we have the inequality

$$|f(A)| \le \xi_0(A) I + \sum_{k=1}^{\infty} \xi_k(A)|W|^k = \sum_{k=0}^{\infty} \xi_k(A)|W|^k. \tag{6.4}$$

Here $|W|^0 = I$.

Let $\|A\|_l$ denote a lattice norm of A. That is, $\|A\|_l \le \||A|\|_l$, and $\|A\|_l \le \|\tilde{A}\|_l$ whenever $0 \le A \le \tilde{A}$. Now the previous theorem implies the inequality

$$\|f(A) - f(S)\|_l \le \sum_{k=1}^{\infty} \xi_k(A) \||W|^k\|_l \tag{6.5}$$

and therefore,

$$\|f(A)\|_l \le \sum_{k=0}^{\infty} \xi_k(A) \||W|^k\|_l.$$

2.7 Diagonalizable matrices

The results of this section are useful for investigation of the forced oscillations of equations close to ordinary ordinary differential equations (see Section 12.3).

2.7.1 A bound for similarity constants of matrices

Everywhere in this section it is assumed that the eigenvalues $\lambda_k = \lambda_k(A)$ $(k = 1, \dots, n)$ of A, taken with their algebraic multiplicities, are geometrically simple. That is, the geometric multiplicity of each eigenvalue is equal to one. As it is well known, in this case A is diagonalizable: there are biorthogonal sequences $\{u_k\}$ and $\{v_k\}$: $(v_j, u_k) = 0 \ (j \neq k)$, $(v_j, u_j) = 1$ $(j, k = 1, \dots, n)$, such that

$$A = \sum_{k=1}^{n} \lambda_k Q_k, \tag{7.1}$$

where $Q_k = (., u_k)v_k$ $(k = 1, \dots, n)$ are one-dimensional eigenprojections. Besides, there is an invertible operator T and a normal operator S, such that

$$TA = ST. \tag{7.2}$$

The constant (the conditional number)

$$\kappa_T := \|T\| \, \|T^{-1}\|$$

is very important for various applications, cf. [103]. That constant is mainly numerically calculated. In the present subsection we suggest a sharp bound for κ_T. Applications of the obtained bound are also discussed.

Denote by $\mu_j, j = 1, \dots, m \leq n$ the distinct eigenvalues of A, and by p_j the algebraic multiplicity of μ_j. In particular, one can write

$$\mu_1 = \lambda_1 = \cdots = \lambda_{p_1}, \ \mu_2 = \lambda_{p_1+1} = \cdots = \lambda_{p_1+p_2},$$

etc.

Let δ_j be the half-distance from μ_j to the other eigenvalues of A:

$$\delta_j := \min_{k=1,\dots,m;\ k\neq j} |\mu_j - \mu_k|/2$$

and

$$\delta(A) := \min_{j=1,\dots,m} \delta_j = \min_{j,k=1,\dots,m;\ k\neq j} |\mu_j - \mu_k|/2.$$

Put

$$\eta\,(A) := \sum_{k=1}^{n-1} \frac{g^k(A)}{\delta^k(A)\sqrt{k!}}.$$

According to (1.1),

$$\eta\,(A) \le \sum_{k=1}^{n-1} \frac{(\sqrt{2}N_2(A_I))^k}{\delta^k(A)\sqrt{k!}}.$$

In [51, Corollary 3.6], the inequality

$$\kappa_T \le \sum_{j=1}^{m} p_j \sum_{k=0}^{n-1} \frac{g^k(A)}{\delta_j^k\sqrt{k!}} \le n(1+\eta\,(A)) \tag{7.3}$$

has been derived. This inequality is not sharp: if A is a normal matrix, then it gives $\kappa_T \le n$ but $\kappa_T = 1$ in this case. In this section we improve inequality (7.3). To this end put

$$\gamma(A) = \begin{cases} n(1+\eta\,(A)) & \text{if } \eta\,(A) \ge 1, \\ (\eta\,(A)+1)[\frac{2\sqrt{n}\eta\,(A)}{1-\eta\,(A)}+1] & \text{if } \eta\,(A) < 1. \end{cases}$$

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

Theorem 2.7.1. *Let A be a diagonalizable $n\times n$-matrix. Then $\kappa_T \le \gamma(A)$.*

The proof of this theorem is presented in the next subsection. Theorem 2.7.1 is sharp: if A is normal, then $g(A) = 0$. Therefore $\eta\,(A) = 0$ and $\gamma(A) = 1$. Thus we obtain the equality $\kappa_T = 1$.

2.7.2 Proof of Theorem 2.7.1

We need the following lemma [51, Lemma 3.4].

Lemma 2.7.2. *Let A be a diagonalizable $n\times n$-matrix and*

$$S = \sum_{k=1}^{n} \lambda_k(.,d_k)d_k, \tag{7.4}$$

where $\{d_k\}$ is an orthonormal basis. Then the operator

$$T = \sum_{k=1}^{n} (.,d_k)v_k \tag{7.5}$$

has the inverse one defined by

$$T^{-1} = \sum_{k=1}^{n} (.,u_k)d_k,$$

and (7.2) holds.

Note that one can take $\|u_k\| = \|v_k\|$. This leads to the equality

$$\|T^{-1}\| = \|T\|. \tag{7.6}$$

We need also the following technical lemma.

Lemma 2.7.3. *Let L_1 and L_2 be projections satisfying the condition $r := \|L_1 - L_2\| < 1$. Then for any eigenvector f_1 of L_1 with $\|f_1\| = 1$ and $L_1 f_1 = f_1$, there exists an eigenvector f_2 of L_2 with $\|f_2\| = 1$ and $L_2 f_2 = f_2$, such that*

$$\|f_1 - f_2\| \leq \frac{2r}{1-r}.$$

Proof. We have $\|L_2 f_1 - L_1 f_1\| \leq r < 1$ and

$$b_0 := \|L_2 f_1\| \geq \|L_1 f_1\| - \|(L_1 - L_2) f_1\| \geq 1 - r > 0.$$

Thanks to the relation $L_2(L_2 f_1) = L_2 f_1$, we can assert that $L_2 f_1$ is an eigenvector of L_2. Then

$$f_2 := \frac{1}{b_0} L_2 f_1$$

is a normed eigenvector of L_2. So

$$f_1 - f_2 = L_1 f_1 - \frac{1}{b_0} L_2 f_1 = f_1 - \frac{1}{b_0} f_1 + \frac{1}{b_0}(L_1 - L_2) f_1.$$

But

$$\frac{1}{b_0} \leq \frac{1}{1-r}$$

and

$$\|f_1 - f_2\| \leq \left(\frac{1}{b_0} - 1\right) \|f_1\| + \frac{1}{b_0} \|(L_1 - L_2) f_1\| \leq \frac{1}{1-r} - 1 + \frac{r}{1-r} = \frac{2r}{1-r},$$

as claimed. $\square$

In the rest of this subsection $d_k = e_k$ $(k = 1, \dots n)$, where $\{e_k\}_{k=1}^n$ is the Schur orthonormal basis of A. Then $S = D$, where D is the diagonal part of A (see Section 2.2). For a fixed index $j \leq m$, let $\hat{P}_j$ be the eigenprojection of D, and $\hat{Q}_j$ the eigenprojection of A corresponding to the same (geometrically simple) eigenvalue μ_j. So

$$A = \sum_{j=1}^{m} \mu_j \hat{Q}_j.$$

The following inequality is proved in [51, inequality (3.3)]:

$$\|\hat{Q}_j - \hat{P}_j\| \leq \eta_{\,j}, \quad \text{where} \quad \eta_{\,j} := \sum_{k=1}^{n-1} \frac{g^k(A)}{\delta_j^k \sqrt{k!}}. \tag{7.7}$$

Let v_{js} and e_{js} $(s = 1, \dots, p_j)$ be the eigenvectors of $\hat{Q}_j$ and $\hat{P}_j$, respectively, and $\|e_{js}\| = 1$. Inequality (7.7) and the previous lemma yield

Corollary 2.7.4. *Assume that*

$$\eta(A) < 1. \tag{7.8}$$

Then

$$\left\| \frac{v_{js}}{\|v_{js}\|} - e_{js} \right\| \le \psi(A) \ \ (s = 1, \dots, p_j), \quad \textit{where} \quad \psi(A) = \frac{2\eta(A)}{1 - \eta(A)}.$$

Hence it follows that

$$\left\| \frac{v_k}{\|v_k\|} - e_k \right\| \le \psi(A) \ \ (k = 1, \dots, n).$$

Taking in the previous corollary Q_j^* instead of Q_j we arrive at the similar inequality

$$\left\| \frac{u_k}{\|u_k\|} - e_k \right\| \le \psi(A) \ \ (k = 1, \dots, n). \tag{7.9}$$

Proof of Theorem 2.7.1. By (7.5) we have

$$\|Tx\| = \left[\sum_{k=1}^n |(x, u_k)|^2\right]^{1/2} = \left[\sum_{k=1}^n |(x, \|u_k\| e_k) + (x, u_k - \|u_k\| e_k)|^2\right]^{1/2}$$
$$\le \left[\sum_{k=1}^n \|u_k\|^2 |(x, \frac{u_k}{\|u_k\|} - e_k)|^2\right]^{1/2} + \left[\sum_{k=1}^n |\|u_k\|(x, e_k)|^2\right]^{1/2} \quad (x \in \mathbb{C}^n).$$

Hence, under condition (7.8), inequality (7.9) implies

$$\|Tx\| \le \max_k \|u_k\| \|x\| (\psi(A)\sqrt{n} + 1) \ \ (x \in \mathbb{C}^n).$$

Thanks to Corollary 3.3 from [51], $\max_k \|u_k\|^2 \le 1 + \eta(A)$. Thus,

$$\|T\|^2 \le \tilde{\gamma}(A) := (\eta(A) + 1) \left[\frac{2\sqrt{n}\eta(A)}{1 - \eta(A)} + 1\right].$$

According to (7.6), $\|T^{-1}\|^2 \le \tilde{\gamma}(A)$. So under condition (7.8), the inequality $\kappa_T \le \tilde{\gamma}(A)$ is true. Combining this inequality with (7.3), we get the required result. □

2.7.3 Applications of Theorem 2.7.1

Theorem 2.7.1 immediately implies

Corollary 2.7.5. *Let A be a diagonalizable $n \times n$-matrix and $f(z)$ be a scalar function defined on the spectrum of A. Then $\|f(A)\| \le \gamma(A) \max_k |f(\lambda_k)|$.*

In particular, we have

$$\|R_z(A)\| \leq \frac{\gamma(A)}{\rho(A,\lambda)} \quad \text{and} \quad \|e^{At}\| \leq \gamma(A)e^{\alpha(A)t} \;\; (t \geq 0).$$

Let A and $\tilde{A}$ be complex $n \times n$-matrices whose eigenvalues λ_k and $\tilde{\lambda}_k$, respectively, are taken with their algebraic multiplicities. Recall that

$$sv_A(\tilde{A}) := \max_k \min_j |\tilde{\lambda}_k - \lambda_j|.$$

Corollary 2.7.6. *Let A be diagonalizable. Then $sv_A(\tilde{A}) \leq \gamma(A)\|A - \tilde{A}\|$.*

Indeed, the operator $S = TAT^{-1}$ is normal. Put $B = T\tilde{A}T^{-1}$. Thanks to the well-known Corollary 3.4 [103], $sv_S(B) \leq \|S - B\|$. Now the required result is due to Theorem 2.7.1.

Furthermore let $\tilde{D}, V_+$ and V_- be the diagonal, upper nilpotent part and lower nilpotent part of matrix $A = (a_{kk})$, respectively. Using the preceding corollary with $A_+ = \tilde{D} + V_+$, we arrive at the relations

$$\sigma(A_+) = \sigma(\tilde{D}), \quad \text{and} \quad \|A - A_+\| = \|V_-\|.$$

Due to the previous corollary we get

Corollary 2.7.7. *Let $A = (a_{jk})_{j,k=1}^n$ be an $n \times n$-matrix, whose diagonal has the property*

$$a_{jj} \neq a_{kk} \;\; (j \neq k;\; k = 1, \ldots, n).$$

Then for any eigenvalue μ of A, there is a $k = 1, \ldots, n$, such that

$$|\mu - a_{kk}| \leq \gamma(A_+)\|V_-\|,$$

and therefore the (upper) spectral radius satisfies the inequality

$$r_s(A) \leq \max_{k=1,\ldots,n} |a_{kk}| + \gamma(A_+)\|V_-\|,$$

and the lower spectral radius satisfies the inequality

$$r_s(A) \geq \min_{k=1,\ldots,n} |a_{kk}| - \gamma(A_+)\|V_-\|,$$

provided $|a_{kk}| > \delta_A$ $(k = 1, \ldots, n)$.

Clearly, one can exchange the places V_+ and V_-.

2.7.4 Additional norm estimates for functions of diagonalizable matrices

Again A is a diagonalizable $n \times n$-matrix, and μ_j $(j = 1, \dots, m \leq n)$ are the distinct eigenvalues of A: $\mu_j \neq \mu_k$ for $k \neq j$ and enumerated in an arbitrary order. Thus, (7.1) can be written as

$$A = \sum_{k=1}^{m} \mu_k \hat{Q}_k,$$

where $\hat{Q}_k, k \leq m$, is the eigenprojection whose dimension is equal to the algebraic multiplicity of μ_k.

Besides, for any function f defined on $\sigma(A)$, $f(A)$ can be represented as

$$f(A) = \sum_{k=1}^{m} f(\mu_k) \hat{Q}_k.$$

Put

$$\tilde{Q}_j = \sum_{k=1}^{j} \hat{Q}_k \ \ (j = 1, \dots, m).$$

Note that according to (7.2) $\tilde{Q}_j = T^{-1} \hat{P}_j T$, where $\hat{P}_j$ is an orthogonal projection. Thus by Theorem 2.7.1,

$$\max_{1 \leq k \leq m} \|\tilde{Q}_k\| \leq \kappa_T \leq \gamma(A).$$

Theorem 2.7.8. *Let A be diagonalizable. Then*

$$\|f(A) - f(\mu_m) I\| \leq \max_{1 \leq k \leq m} \|\tilde{Q}_k\| \sum_{k=1}^{m-1} |f(\mu_k) - f(\mu_{k+1})|.$$

To prove this theorem, we need the following analog of the Abel transform.

Lemma 2.7.9. *Let a_k be numbers and W_k $(k = 1, \dots, m)$ be bounded linear operators in a Banach space. Let*

$$\Psi := \sum_{k=1}^{m} a_k W_k \quad \textit{and} \quad B_j = \sum_{k=1}^{j} W_k.$$

Then

$$\Psi = \sum_{k=1}^{m-1} (a_k - a_{k+1}) B_k + a_m B_m.$$

Proof. Obviously,

$$\Psi = a_1 B_1 + \sum_{k=2}^{m} a_k (B_k - B_{k-1}) = \sum_{k=1}^{m} a_k B_k - \sum_{k=2}^{m} a_k B_{k-1} = \sum_{k=1}^{m} a_k B_k - \sum_{k=1}^{m-1} a_{k+1} B_k,$$

as claimed. □

Proof of Theorem 2.7.8. According to Lemma 2.7.9,

$$f(A) = \sum_{k=1}^{m-1} (f(\mu_k) - f(\mu_{k+1}))\tilde{Q}_k + f(\mu_m)\tilde{Q}_m. \tag{7.10}$$

But $\tilde{Q}_m = I$. Hence, the assertion of the theorem at once follows. □

According to the previous theorem we have

Corollary 2.7.10. *Assume that $f(\lambda)$ is real for all $\lambda \in \sigma(A)$, and that the condition*

$$f(\mu_{k+1}) \le f(\mu_k) \ \ (k = 1, \dots, m-1) \tag{7.11}$$

holds. Then

$$\|f(A) - f(\mu_m) I\| \le \max_{1 \le k \le m} \|\tilde{Q}_k\| [f(\mu_1) - f(\mu_m)].$$

From this corollary it follows that

$$\|f(A)\| \le \max_{1 \le k \le m} \|\tilde{Q}_k\| [f(\mu_1) + |f(\mu_m)| - f(\mu_m)],$$

provided (7.11) holds.

2.8 Matrix exponential

2.8.1 Lower bounds

As it was shown in Section 2.5,

$$\|\exp(At)\| \le e^{\alpha(A)t} \sum_{k=0}^{n-1} g^k(A) t^k \frac{\gamma_{n,k}}{k!} \quad (t \ge 0)$$

where $\alpha(A) = \max_{k=1,\dots,n} \operatorname{Re} \lambda_k(A)$. Moreover, by (3.1),

$$\|\exp(At)\| \le e^{\alpha(A)t} \sum_{k=0}^{n-1} \frac{g^k(A) t^k}{(k!)^{3/2}} \quad (t \ge 0). \tag{8.1}$$

Taking into account that the operator $\exp(-At)$ is the inverse one to $\exp(At)$ it is not hard to show that

$$\|\exp(At)h\| \geq \frac{e^{\beta(A)t}\|h\|}{\sum_{k=0}^{n-1} g^k(A)t^k(k!)^{-1}\gamma_{n,k}} \quad (t \geq 0, h \in \mathbb{C}^n),$$

where $\beta(A) = \min_{k=1,\dots,n} \operatorname{Re}\lambda_k(A)$. Therefore by (3.1),

$$\|\exp(At)h\| \geq \frac{e^{\beta(A)t}\|h\|}{\sum_{k=0}^{n-1} g^k(A)(k!)^{-3/2}\ t^k} \quad (t \geq 0). \tag{8.2}$$

Moreover, if A is a diagonalizable $n \times n$-matrix, then due to Theorem 2.7.1 we conclude that

$$\|e^{-At}\| \leq \gamma(A)e^{-\beta(A)t} \ \ (t \geq 0).$$

Hence,

$$\|e^{At}h\| \geq \frac{\|h\|e^{\beta(A)t}}{\gamma(A)} \quad (t \geq 0, h \in \mathbb{C}^n).$$

2.8.2 Perturbations of matrix exponentials

Let $A, \tilde{A} \in \mathbb{C}^{n\times n}$ and $E = \tilde{A} - A$. To investigate perturbations of matrix exponentials one can use the equation

$$e^{\tilde{A}t} - e^{At} = \int_0^t e^{\tilde{A}(t-s)}Ee^{As}ds \tag{8.3}$$

and estimate (8.1). Here we investigate perturbations in the case when $\|\tilde{A}E - EA\|$ is small enough. We will say that A is stable (Hurwitzian), if $\alpha(A) < 0$. Assume that A is stable and put

$$u(A) = \int_0^\infty \|e^{At}\|dt \quad \text{and} \quad v_A = \int_0^\infty t\|e^{At}\|dt.$$

Theorem 2.8.1. *Let A be stable, and*

$$\|\tilde{A}E - EA\|v_A < 1. \tag{8.4}$$

Then $\tilde{A}$ is also stable. Moreover,

$$u(\tilde{A}) \leq \frac{u(A) + v_A\|E\|}{1 - v_A\|\tilde{A}E - EA\|} \tag{8.5}$$

and

$$\int_0^\infty \|e^{\tilde{A}t} - e^{At}\|dt \leq \|E\|v_A + \frac{\|\tilde{A}E - EA\|v_A(u(A) + v_A\|E\|)}{1 - v_A\|\tilde{A}E - EA\|}. \tag{8.6}$$

This theorem is proved in the next subsection.

Furthermore, by (8.1) we obtain $u(A) \le u_0(A)$ and $v_A \le \hat{v}_A$, where

$$u_0(A) := \sum_{k=0}^{n-1} \frac{g^k(A)}{|\alpha(A)|^{k+1}(k!)^{1/2}} \quad \text{and} \quad \hat{v}_A := \sum_{k=0}^{n-1} \frac{(k+1)g^k(A)}{|\alpha(A)|^{k+2}(k!)^{1/2}}.$$

Thus, Theorem 2.8.1 implies

Corollary 2.8.2. *Let A be stable and $\|\tilde{A}E - EA\|\hat{v}_A < 1$. Then $\tilde{A}$ is also stable. Moreover,*

$$u(\tilde{A}) \le \frac{u_0(A) + \hat{v}_A\|E\|}{1 - \hat{v}_A\|\tilde{A}E - EA\|}$$

and

$$\int_0^\infty \|e^{\tilde{A}t} - e^{At}\|dt \le \|E\|\hat{v}_A + \frac{\|\tilde{A}E - EA\|\hat{v}_A(u_0(A) + \hat{v}_A\|E\|)}{1 - \hat{v}_A\|\tilde{A}E - EA\|}.$$

2.8.3 Proof of Theorem 2.8.1

We use the following result, let $f(t), c(t)$ and $h(t)$ be matrix functions defined on $[0, b]$ $(0 < b < \infty)$. Besides, f and h are differentiable and c integrable. Then

$$\int_0^b f(t)c(t)h(t)dt = f(b)j(b)h(b) - \int_0^b (f'(t)j(t)h(t) + f(t)j(t)h'(t))dt$$

with

$$j(t) = \int_0^t c(s)ds.$$

For the proof see Lemma 8.2.1 below. By that result

$$e^{\tilde{A}t} - e^{At} = \int_0^t e^{A(t-s)}Ee^{As}ds = Ete^{At} + \int_0^t e^{\tilde{A}(t-s)}[\tilde{A}E - EA]se^{As}ds.$$

Hence,

$$\int_0^\infty \|e^{\tilde{A}t} - e^{At}\|dt \le \int_0^\infty \|Ete^{At}\|dt + \int_0^\infty \int_0^t \|e^{\tilde{A}(t-s)}\|\|\tilde{A}E - EA\|\|se^{As}\|ds\, dt.$$

But

$$\begin{aligned} \int_0^\infty \int_0^t \|e^{\tilde{A}(t-s)}\|s\|e^{As}\|ds\, dt &= \int_0^\infty \int_s^\infty \|e^{\tilde{A}(t-s)}\|s\|e^{As}\|dt\, ds \\ &= \int_0^\infty s\|e^{As}\|ds \int_0^\infty \|e^{\tilde{A}t}\|dt = v_A u(\tilde{A}). \end{aligned}$$

Thus

$$\int_0^\infty \|e^{\tilde{A}t} - e^{At}\| dt \le \|E\| v_A + \|\tilde{A}E - EA\| v_A u(\tilde{A}). \tag{8.7}$$

Hence,

$$u(\tilde{A}) \le u(A) + \|E\| v_A + \|\tilde{A}E - EA\| v_A u(\tilde{A}).$$

So according to (8.4), we get (8.5). Furthermore, due to (8.7) and (8.5) we get (8.6) as claimed. □

2.9 Matrices with non-negative off-diagonals

In this section it is assumed that $A = (a_{ij})_{j,k=1}^n$ is a real matrix with

$$a_{ij} \ge 0 \text{ for } i \ne j. \tag{9.1}$$

Put

$$a = \min_{j=1,\dots,n} a_{jj} \quad \text{and} \quad b = \max_{j=1,\dots,n} a_{jj}.$$

For a scalar function $f(\lambda)$ write

$$\alpha_k(f, A) := \inf_{a \le x \le b} \frac{f^{(k)}(x)}{k!} \quad \text{and} \quad \beta_k(f, A) := \sup_{a \le x \le b} \frac{f^{(k)}(x)}{k!} \quad (k = 0, 1, 2, \dots),$$

assuming that the derivatives exist.

Let $W = A - \operatorname{diag}(a_{jj})$ be the off diagonal part of A.

Theorem 2.9.1. *Let condition* (9.1) *hold and $f(\lambda)$ be holomorphic on a neighborhood of a Jordan set, whose boundary C has the property*

$$|z - a_{jj}| > \sum_{k=1, k \ne j}^n a_{jk}$$

for all $z \in C$ and $j = 1, \dots, n$. In addition, let f be real on $[a, b]$. Then the following inequalities are valid:

$$f(A) \ge \sum_{k=0}^\infty \alpha_k(f, A) W^k, \tag{9.2}$$

provided

$$r_s(W) \lim_{k \to \infty} \sqrt[k]{|\alpha_k(f, A)|} < 1,$$

and

$$f(A) \le \sum_{k=0}^\infty \beta_k(f, A) W^k, \tag{9.3}$$

provided,

$$r_s(W) \lim_{k \to \infty} \sqrt[k]{|\beta_k(f, A)|} < 1.$$

In particular, if $\alpha_k(f, A) \ge 0$ $(k = 0, 1, \dots)$, then matrix $f(A)$ has non-negative entries.

Proof. By (6.2) and (6.3),

$$f(A) = f(S) + \sum_{k=1}^{\infty} M_k,$$

where

$$M_k = \sum_{j_1=1}^{n} \hat{Q}_{j_1} W \sum_{j_2=1}^{n} \hat{Q}_{j_2} W \dots W \sum_{j_{k+1}=1}^{n} \hat{Q}_{j_{k+1}} J_{j_1 j_2 \dots j_{k+1}}.$$

Here

$$J_{j_1 \dots j_{k+1}} = \frac{(-1)^{k+1}}{2\pi i} \int_C \frac{f(\lambda) d\lambda}{(b_{j_1} - \lambda) \dots (b_{j_{k+1}} - \lambda)} \quad (b_j = a_{jj}).$$

Since S is real, Lemma 1.5.2 from [31] gives us the inequalities

$$\alpha_k(f, A) \le J_{j_1 \dots j_{k+1}} \le \beta_k(f, A).$$

Hence,

$$M_k \ge \alpha_k(f, A) \sum_{j_1=1}^{n} \hat{Q}_{j_1} W \sum_{j_2=1}^{n} \hat{Q}_{j_2} W \dots W \sum_{j_{k+1}=1}^{n} \hat{Q}_{j_{k+1}} = \alpha_k(f, A) W^k.$$

Similarly, $M_k \le \beta_k(f, A) W^k$. This implies the required result. □

2.10 Comments

One of the first estimates for the norm of a regular matrix-valued function was established by I.M. Gel'fand and G.E. Shilov [19] in connection with their investigations of partial differential equations, but that estimate is not sharp; it is not attained for any matrix. The problem of obtaining a sharp estimate for the norm of a matrix-valued function has been repeatedly discussed in the literature, cf. [14]. In the late 1970s, the author obtained a sharp estimate for a matrix-valued function regular on the convex hull of the spectrum, cf. [21] and references therein. It is attained in the case of normal matrices. Later, that estimate was extended to various classes of non-selfadjoint operators, such as Hilbert–Schmidt operators, quasi-Hermitian operators (i.e., linear operators with completely continuous imaginary components), quasiunitary operators (i.e., operators represented as a sum of a unitary operator and a compact one), etc. For more details see [31, 56] and references given therein.

The material of this chapter is taken from the papers [51, 53, 56] and the monograph [31].

About the relevant results on matrix-valued functions and perturbations of matrices see the well-known books [9, 74] and [90].

http://www.springer.com/978-3-0348-0576-6

Stability of Vector Differential Delay Equations
Gil', M.I.
2013, X, 259 p., Softcover
ISBN: 978-3-0348-0576-6
A product of Birkhäuser Basel