Integr. equ. oper. theory 30 (1998) 326 - 337
0378-620X/98/030326-12 $ 1.50+0.20/0

Integral Equations
and Operator Theory

ON SOME PROPERTIES OF FACTORIZATION INDICES[1]

Israel Feldman and Alexander Markus

Dedicated to the memory of Mark Grigorievich Krein

Various kinds of factorization indices are considered (right partial indices, left partial indices, Birkhoff indices), and some connections between them are described. We solve also the problem on the relation between the partial indices of two matrix functions and of their product.

1. Introduction

Factorization of matrix functions and corresponding factorization indices (or partial indices) have various applications in analysis. We mention here singular integral equations, Wiener-Hopf equations, Toeplitz operators, finite section method for convolution equations (see, e.g., [3, 5, 8, 10]).

There are two known types of factorizations of matrix functions – right and left factorizations. The formal difference between them is non-essential but they have different applications. G. D. Birkhoff [1] introduced one more type of factorization and corresponding factorization indices.

The principal goal of this paper is to study relations between various types of factorization indices. The main result consists in complete description of the connection between the right indices and the Birkhoff indices (Section 3). This description is given in terms of majorization in the sense of Hardy, Littlewood and Polya [9]. Regarding the right and left indices, we prove that the unique relation between them is the equality of their sums (Theorem 4.1).

We also solve the problem on the connection between the right indices of two matrix functions and of their product. It turns out that the obvious necessary condition is also sufficient (Theorem 4.3).

It is known [1, 3] that the Birkhoff indices are not determined uniquely. In Section 5 we discuss the question about the number of different collections of the Birkhoff indices.

Factorization of matrix functions and its applications is one of the numerous areas of analysis which was developed under strong influence of M. G. Krein. We mention here the

[1]This research was supported by the Israel Science Foundation founded by the Israel Academy of Sciences and Humanities.

classical work by Gohberg and Krein [6]. One of results of this work (see Theorem 2.6 below) will be systematically used in our work.

We are kindly grateful to I. Gohberg for helpful discussions.

2. Preliminaries

2.1. This section contains the main definitions and known results we need. For the sake of simplicity we consider only rational matrix functions and only the unit circle $\mathbf{T}=\{t : |t| = 1\}$.

Let $\mathcal{R}^{n\times n}$ denote the algebra of all rational $n \times n$ matrix functions which have no poles on $\mathbf{T}$. By $\mathcal{R}_+^{n\times n}(\mathcal{R}_-^{n\times n})$ we denote its subalgebras consisting of all matrix functions which have no poles in $|t| \leq 1$ (in $|t| \geq 1$, respectively). Denote by $G\mathcal{R}^{n\times n}$, $G\mathcal{R}_+^{n\times n}$, $G\mathcal{R}_-^{n\times n}$ the sets of invertible elements in corresponding algebras.

Let $A(t) \in G\mathcal{R}^{n\times n}$. The *right factorization* of $A(t)$ (with respect to $\mathbf{T}$) is its representation in the form

$$A(t) = A_-(t)D_r(t)A_+(t)$$

where

$$A_+(t) \in G\mathcal{R}_+^{n\times n}, \quad A_-(t) \in G\mathcal{R}_-^{n\times n}, \quad D_r(t) = \mathrm{diag}[t^{r_i}]_1^n \quad (r = (r_j)_1^n \in \mathbf{Z}^n).$$

The integers $(r_j)_1^n$ are called the *right indices* of $A(t)$.

It is well known [3, Theorems 1.1.1 and 1.2.1] that the right factorization always exists and the right indices are determined uniquely up to their order. It is easy to see that arbitrary permutation of right indices is possible. Let us note also, that

$$\sum_1^n r_j = \mathrm{ind}\ \det A(t), \tag{2.1}$$

where ind det $A(t)$ is the index of the function $\det A(t)$ with respect $\mathbf{T}$.

The *left factorization* and the *left indices* of $A(t)$ are defined by the equalities

$$A(t) = B_+(t)D_l(t)B_-(t), \quad l = (l_j)_1^n \in \mathbf{Z}^n,$$

where $B_\pm(t) \in G\mathcal{R}_\pm^{n\times n}$. Obviously,

$$\sum_1^n l_j = \mathrm{ind}\ \det A(t). \tag{2.2}$$

The first known property of factorization indices we need is the following (see, e.g., proof of Theorem 1.2.1 in [3]).

Lemma 2.1. *If $A(t) \in G\mathcal{R}^{n\times n} \cap \mathcal{R}_+^{n\times n}$, then all right indices of $A(t)$ are non-negative.*

2.2. Let $L_2^n(\mathbf{T})$ and H_2^n be the spaces of vector functions with components from $L_2(\mathbf{T})$ and H_2, respectively. For a continuous matrix function $A(t)$ $(t \in\mathbf{T})$ we define the corresponding *Toeplitz operator* T_A on the space H_2^n by the equality

$$T_A f = P(A(t)f(t)).$$

Here P is the orthogonal projection in $L_2^n(\mathrm{T})$ on H_2^n.

We will use the following simple and well-known property of Toeplitz operators .

Lemma 2.2. *If $A(t) \in \mathcal{R}^{n\times n}$ and $A_+(t) \in \mathcal{R}_+^{n\times n}$, then $T_{AA_+} = T_A T_{A_+}$.*

Let S be a bounded linear operator in a Hilbert space H. Denote by $\operatorname{Im} S$ the range of S, by $\ker S$ its kernel and by $\operatorname{coker} T$ the orthogonal complement of $\operatorname{Im} T$ in H. The following result (see, e.g., [3, p. 74]) shows the role of right indices in the theory of Toeplitz operators.

Theorem 2.3. *If $A(t) \in G\mathcal{R}^{n\times n}$, then $\operatorname{Im} T_A$ is closed and*

$$\dim \ker T_A = -\sum_1^n r_j^-, \quad \dim \operatorname{coker} T_A = \sum_1^n r_j^+.$$

Here and in what follows we use the standart notations:

$$r^+ = \max(r, 0), \quad r^- = \min(r, 0) \quad (r \in \mathbf{R}).$$

For a continuous $n \times n$ matrix function $A(t)$ $(t \in \mathrm{T})$ we denote by $\|A(t)\|$ the norm of the operator $A(t)$ in the n-dimensional Hilbert space $\mathbf{C}^n$. Set $\|A\|_\infty = \max\{\|A(t)\| : t \in \mathrm{T}\}$.

Lemma 2.4 *If $A(t) \in \mathcal{R}^{n\times n}$ and $\|A\|_\infty < 1$, then all right and left indices of the matrix function $I + A(t)$ are equal to zero.*

Proof of this lemma may be found, e.g., in [3, p. 75].

2.3. For a vector $a = (a_j)_1^n \in \mathbf{R}^n$ we denote by $a^* = (a_j^*)_1^n$ its decreasing rearrangement: $a_1^* \geq a_1^* \geq \cdots \geq a_n^*$. The following definition belongs to Hardy, Littlewood and Polya [9].

Let $a,\ b \in \mathbf{R}^n$. We say that the vector a is *majorized* by the vector b $(a \prec b)$ if

$$\sum_1^n a_j = \sum_1^n b_j \quad \text{and} \quad \sum_1^k a_j^* \leq \sum_1^k b_j^* \quad (k = 1,\ 2, \ldots,\ n-1).$$

This important notion has many applications in various areas of mathematics (see [11]).

Let us begin with an equivalent property. We consider here and in what follows the notion of majorization only for vectors with integer coordinates.

Lemma 2.5. *For two vectors $a,\ b \in \mathbf{Z}^n$ the following properties are equivalent:*

(1) $a \prec b$;

(2) $\sum_{j=1}^n a_j = \sum_{j=1}^n b_j$ and $\sum_{j=1}^n (a_j - m)^+ \leq \sum_{j=1}^n (b_j - m)^+$ for all $m \in \mathbf{Z}$.

Analogous statement with $\mathbf{R}$ instead of $\mathbf{Z}$ is proved in [11, p. 109]. The same proof is suitable for Lemma 2.5.

2.4. The following result of I. Gohberg and M. Krein [6] (see also [10, Theorem 6.5]) plays an important role in this paper.

Theorem 2.6. *Let $A(t) \in G\mathcal{R}^{n\times n}$, and let r be collection of its right indices. If $s \in \mathbf{Z}^n$ and $s \prec r$ then for any $\varepsilon > 0$ there exists a matrix function $A_\varepsilon(t) \in \mathcal{R}^{n\times n}$ such that $\|A_\varepsilon - A\|_\infty < \varepsilon$ and the collection of right indices of $A_\varepsilon(t)$ is s.*

Note that in [6] the inverse result was also obtained, but we will not use it. Another remark is that in [6] both $A(t)$ and $A_\varepsilon(t)$ belong to some wider class of matrix functions. It is easy to see by inspecting the proof from [6] that in the case $A(t) \in \mathcal{R}^{n\times n}$ the matrix function $A_\varepsilon(t)$ can also be chosen from $\mathcal{R}^{n\times n}$.

2.5. In connection with some problems of ordinary differential equations Birkhoff [1] considered factorization of the following form:

$$A(t) = F_-(t)F_+(t)D_b(t) \quad (b = (b_j)_1^n \in \mathbf{Z}^n). \tag{2.3}$$

He proved the existence of this factorization for rather restricted class of matrix functions $A(t)$. Under natural assumptions this result was obtained in [2] (see also [3], Theorem 6.2.2). It follows from the proof in [3] that for $A(t) \in G\mathcal{R}^{n\times n}$ the factors $F_\pm(t)$ belong to $G\mathcal{R}_\pm^{n\times n}$, respectively. The numbers b_j $(j = 1, 2, \ldots, n)$ in (2.3) will be called the *Birkhoff indices* of $A(t)$. Obviously,

$$\sum_1^n b_j = \operatorname{ind}\det A(t). \tag{2.4}$$

Birkhoff [1] mentioned that the indices $(b_j)_1^n$ are not determined uniquely. Such an example is contained in [3, p. 133]. In this example there are exactly two different collections of Birkhoff indices. In the last section of this paper we will show that the number of different collections of these indices may be arbitrary. On the other hand, in many cases (e.g., for $A(t) = D_b(t)$) the Birkhoff indices are uniquely determined (even the permutations of the numbers $(b_j)_1^n$ are not possible).

3. The right indices and the Birkhoff indices

3.1. **Lemma 3.1.** *If*

$$B_+(t)D_b(t) = A_-(t)D_r(t)A_+(t), \tag{3.1}$$

where $A_+(t)$, $B_+(t) \in G\mathcal{R}_+^{n\times n}$, $A_-(t) \in G\mathcal{R}_-^{n\times n}$, then

$$\min\{b_i : 1 \le i \le n\} \le r_j \le \max\{b_i : 1 \le i \le n\} \quad (j = 1, 2, \ldots, n). \tag{3.2}$$

Proof. According to the notations of Subsection 2.4 we set $b_n^* = \min b_i$. Divide (3.1) by $t^{b_n^*}$:

$$B_+(t)\operatorname{diag}[t^{b_j - b_n^*}]_{j=1}^n = A_-(t)\operatorname{diag}[t^{r_j - b_n^*}]_{j=1}^n A_+(t).$$

The matrix function in the left-hand side of the last equality belongs to $\mathcal{R}_+^{n\times n}$, and by Lemma 2.1 $r_j - b_n^* \ge 0$ $(j = 1, 2, \ldots, n)$. We proved the left part of inequality (3.2). Passing to the transposed and inverse matrices in (3.1) we obtain the right part of inequality (3.2). □

Theorem 3.2. *If $r = (r_j)_1^n$ is the collection of right indices of a matrix function $A(t) \in G\mathcal{R}^{n\times n}$ and $b = (b_j)_1^n$ is one of the collections of its Birkhoff indices, then $r \prec b$.*

Proof. It follows from (2.1) and (2.4) that

$$\sum_1^n r_j = \sum_1^n b_j,$$

and by Lemma 2.5 it is sufficient to prove that for all $m \in \mathbf{Z}$

$$\sum_{j=1}^n (r_j - m)^+ \leq \sum_{j=1}^n (b_j - m)^+. \tag{3.3}$$

If we prove (3.3) for $m = 0$, then we will immediately obtain (3.3) for arbitrary m by considering the matrix function $t^{-m}A(t)$.Hence we need only to prove that

$$\sum_{j=1}^n r_j^+ \leq \sum_{j=1}^n b_j^+. \tag{3.4}$$

By the conditions of the theorem we have

$$F_-(t)F_+(t)D_b(t) = E_-(t)D_r(t)E_+(t)$$

for some $E_+,\ F_+ \in G\mathcal{R}_+^{n\times n}$ and $E_-,\ F_- \in G\mathcal{R}_-^{n\times n}$. Therefore

$$F_+(t)D_b(t) = F_-^{-1}(t)E_-(t)D_r(t)E_+(t). \tag{3.5}$$

Represent the left-hand side of (3.5) in the form

$$F_+(t)D_b(t) = F_+(t)D_b^-(t)D_b^+(t), \tag{3.6}$$

where

$$D_b^{\pm}(t) = \operatorname{diag}\, [t^{b_j^{\pm}}]_{j=1}^n.$$

It follows from (3.6) and from Lemma 2.2 that

$$T_{F_+D_b} = T_{F_+D_b^-}T_{D_b^+}. \tag{3.7}$$

This implies

$$\dim\operatorname{coker} T_{F_+D_b} \leq \dim\operatorname{coker} T_{F_+D_b^-} + \dim\operatorname{coker} T_{D_b^+}. \tag{3.8}$$

Obviously (and also follows from Theorem 2.3) that

$$\dim\operatorname{coker} T_{D_b^+} = \sum_{j=1}^n b_j^+. \tag{3.9}$$

From Lemma 3.1 (right part of inequality (3.2)) it follows that all right indices of the matrix function $F_+(t)D_b^-(t)$ are non-positive. By Theorem 2.3 this implies

$$\dim\ \operatorname{coker} T_{F_+D_b^-} = 0. \tag{3.10}$$

From (3.8) - (3.10) we get

$$\dim\operatorname{coker} T_{F_+D_b} \leq \sum_{j=1}^n b_j^+. \tag{3.11}$$

On the other hand, it follows from (3.5) and from Theorem 2.3 that

$$\dim \operatorname{coker} T_{F_+D_b} = \sum_{j=1}^{n} r_j^+. \tag{3.12}$$

Finally, (3.11) and (3.12) imply (3.4). □

Remark. Theorem 3.2 remains true under more general assumptions (e.g., for any continuous and non-degenerate on **T** matrix function $A(t)$).

3.2. Here we prove the inverse theorem.

Theorem 3.3. *Let $r,\ b \in \mathbf{Z}^n$ and $r \prec b$. Then there exists a matrix function $A(t) \in G\mathcal{R}^{n\times n}$ such that the collection of the right indices of $A(t)$ is r, and one of its collections of Birkhoff indices is b.*

Proof. It follows from the condition $r \prec b$ and from Theorem 2.6 that there exists a matrix function $M(t) \in \mathcal{R}^{n\times n}$ such that $\|M\|_\infty < 1$ and the collection of the right indices of the matrix function $A(t) = D_b(t) + M(t)$ coincides with r. This means that

$$D_b(t) + M(t) = E_-(t)D_r(t)E_+(t), \tag{3.13}$$

where $E_\pm(t) \in G\mathcal{R}_\pm^{n\times n}$. We can write the left-hand side of (3.13) in the form

$$D_b(t) + M(t) = (I + M(t)D_b^{-1}(t))D_b(t). \tag{3.14}$$

From inequality

$$\|MD_b^{-1}\|_\infty \le \|M\|_\infty \|D_b^{-1}\|_\infty = \|M\|_\infty < 1$$

and from Lemma 2.4 we obtain that all right indices of the matrix function $I + M(t)D_b^{-1}(t)$ are equal to 0. In other words,

$$I + M(t)D_b^{-1}(t) = F_-(t)F_+(t), \tag{3.15}$$

where $F_\pm(t) \in G\mathcal{R}_\pm^{n\times n}$. The equalities (3.13) – (3.15) imply

$$A(t) = E_-(t)D_r(t)E_+(t) = F_-(t)F_+(t)D_b(t),$$

which completes the proof. □

Corollary 3.4. *If $r \prec b$ then there exist matrix functions $E_+(t),\ F_+(t) \in G\mathcal{R}_+^{n\times n}$ and $E_-(t) \in G\mathcal{R}_-^{n\times n}$ such that $F_+(t)D_b(t) = E_-(t)D_r(t)E_+(t)$.*

4. The right and left indices. The indices of the product

4.1. If $r = (r_j)_1^n$ and $l = (l_j)_1^n$ are the collections of right and left indices, respectively, of a matrix function $A(t) \in G\mathcal{R}^{n\times n}$, then by (2.1) and (2.2)

$$\sum_{j=1}^{n} r_j = \sum_{j=1}^{n} l_j. \tag{4.1}$$

Here we prove the inverse statement.

Theorem 4.1. *If for two vectors r, l $(\in \mathbf{Z}^n)$ condition (4.1) holds, then there exists a matrix function $A(t) \in \mathcal{GR}^{n\times n}$ such that the collection of its right indices coincides with r and the collection of its left indices coincides with l.*

Proof. Let $s \in \mathbf{Z}^n$ be a vector such that $r \prec s$ and $l \prec s$. For example, we can choose the vector s in the following way:

$$s_1 = \max\left(\sum_{j=1}^n r_j^+,\ \sum_{j=1}^n l_j^+\right), \quad s_2 = s_3 = \cdots = s_{n-1} = 0, \quad s_n = \sum_{j=1}^n r_j - s_1.$$

By corollary 3.4 there exist matrix functions $A_+(t)$, $B_+(t) \in \mathcal{GR}_+^{n\times n}$ and $B_-(t) \in \mathcal{GR}_-^{n\times n}$ such that

$$A_+(t)D_s(t) = B_-(t)D_r(t)B_+(t), \tag{4.2}$$

and also there exist matrix functions $E_+(t)$, $C_+(t) \in \mathcal{GR}_+^{n\times n}$ and $C_-(t) \in \mathcal{GR}_-^{n\times n}$ such that

$$E_+(t)D_s(t) = C_-(t)D_l(t)C_+(t). \tag{4.3}$$

Transpose the equality (4.3):

$$D_s(t)E_+^\top(t) = C_+^\top(t)D_l(t)C_-^\top(t), \tag{4.4}$$

multiply (4.4) from the left by $A_+(t)$ and use the equality (4.2):

$$B_-(t)D_r(t)B_+(t)E_+^\top(t) = A_+(t)C_+^\top(t)D_l(t)C_-^\top(t).$$

This equality proves the theorem. □

Let us remark that Theorem 4.1 was proved in an unpublished work of I. Spitkovsky and Y. Zucker for two particular cases: (1) $n = 2$; (2) either the collection r or the collection l is a *stable* collection (i.e. either $|r_j - r_k| \leq 1$ or $|l_j - l_k| \leq 1$ for all j and k).

Corollary 4.2. *If for two vectors b, $l \in (\mathbf{Z}^n)$ condition*

$$\sum_{j=1}^n b_j = \sum_{j=1}^n l_j$$

holds, then there exists a matrix function $A(t) \in \mathcal{GR}^{n\times n}$ such that the collection of the left indices of $A(t)$ is l, and one of its collections of Birkhoff indices is b.

Proof. By Theorem 4.1 there exist matrix functions $A_+(t)$, $B_+(t) \in \mathcal{GR}_+^{n\times n}$ and $A_-(t)$, $B_-(t) \in \mathcal{GR}_-^{n\times n}$ such that

$$A_-(t)D_b(t)A_+(t) = B_+(t)D_l(t)B_-(t).$$

It follows from [3], Theorem 6.2.2, that there exist matrix functions $C_\pm(t) \in \mathcal{GR}_\pm^{n\times n}$ and a permutation π of the set $\{1, 2, \ldots, n\}$ such that

$$A_-(t)D_b(t)A_+(t) = C_-(t)C_+(t)D_{b'}(t)$$

where $b'_k = b_{\pi(k)}$ $(k = 1, 2, \dots, n)$. Then $D_b(t) = P^{-1} D_{b'}(t) P$ for an appropriate permutation matrix P, and we set

$$A(t) = C_-(t) C_+(t) P D_b(t) = B_+(t) D_l(t) B_-(t) P.$$

□

4.2. Let $A(t),\ B(t) \in G\mathcal{R}^{n\times n}$ and $(p_j)_1^n,\ (q_j)_1^n,\ (r_j)_1^n$ be the right indices of the matrix functions $A(t),\ B(t),\ A(t)B(t)$, respectively. It follows from (2.1) that

$$\sum_{j=1}^n p_j + \sum_{j=1}^n q_j = \sum_{j=1}^n r_j. \tag{4.5}$$

Now the question arises: what additional relations exist between the numbers p_j, q_j and r_j? This question looks like the known problem about relations between the partial multiplicities of two matrix polynomials and their product (see e.g. [7, Corollary 5.2.4]). It turns out that in contrast with the last problem there is no any additional relation between the numbers p_j, q_j and r_j besides (4.5).

Theorem 4.3. *If $(p_j)_1^n,\ (q_j)_1^n,\ (r_j)_1^n$ are integers and the equality (4.5) holds, then there exist matrix functions $A(t),\ B(t) \in G\mathcal{R}^{n\times n}$ such that the right indices of $A(t),\ B(t)$ and $A(t)B(t)$ are $(p_j)_1^n,\ (q_j)_1^n$ and $(r_j)_1^n$, respectively.*

Proof. Choose vectors $a,\ b \in \mathbf{Z}^n$ such that $p \prec a,\ q \prec b$ and $r \prec a+b$. For example, we can set:

$$a_1 = \sum_{j=1}^n p_j^+,\quad b_1 = \max\left(\sum_{j=1}^n q_j^+,\ \sum_{j=1}^n r_j^+ - \sum_{j=1}^n p_j^+\right),\quad a_j = b_j = 0\ \ (j = 2, 3, \dots,\ n-1),$$

$$a_n = \sum_{j=1}^n p_j - a_1,\quad b_n = \sum_{j=1}^n q_j - b_1.$$

It follows from the relation $p \prec a$ and from Theorem 2.6 that for any $\varepsilon > 0$ there exists a matrix function $M_1(t) \in \mathcal{R}^{n\times n}$ such that $\|M_1\|_\infty < \varepsilon$ and

$$D_a(t) + M_1(t) = S_-(t) D_p(t) S_+(t), \tag{4.6}$$

where $S_\pm(t) \in G\mathcal{R}_\pm^{n\times n}$. Similarly,

$$D_b(t) + M_2(t) = R_-(t) D_q(t) R_+(t), \tag{4.7}$$

$$D_{a+b}(t) + M_3(t) = F_-(t) D_r(t) F_+(t), \tag{4.8}$$

where $M_2,\ M_3 \in \mathcal{R}^{n\times n},\ \ R_+,\ F_+ \in G\mathcal{R}_+^{n\times n},\ \ R_-,\ F_- \in G\mathcal{R}_-^{n\times n}$ and $\|M_2\|_\infty < \varepsilon,\ \ \|M_3\|_\infty < \varepsilon$. Rewrite (4.6) – (4.8) in the following form:

$$D_a(I + D_a^{-1} M_1) = S_- D_p S_+, \tag{4.9}$$

$$(I + M_2 D_b^{-1}) D_b = R_- D_q R_+, \tag{4.10}$$

$$D_a(I + D_a^{-1} M_3 D_b^{-1}) D_b = F_- D_r F_+. \tag{4.11}$$

If we find D_a and D_b from (4.9), (4.10) and substitute them in (4.11), we obtain

$$S_- D_p S_+ (I + M) R_- D_q R_+ = F_- D_r F_+, \tag{4.12}$$

where

$$I + M = (I + D_a^{-1} M_1)^{-1} (I + D_a^{-1} M_3 D_b^{-1})(I + M_2 D_b^{-1})^{-1}.$$

For sufficiently small ε we have $\|M\|_\infty < 1$, and by Lemma 2.4 all left indices of $I + M(t)$ are equal to 0, i.e.

$$I + M(t) = E_+(t) E_-(t) \quad (E_\pm \in G\mathcal{R}_\pm^{n\times n}). \tag{4.13}$$

Denote

$$A(t) = S_-(t) D_p(t) S_+(t) E_+(t), \quad B(t) = E_-(t) R_-(t) D_q(t) R_+(t).$$

From (4.12) and (4.13) we obtain

$$A(t) B(t) = F_-(t) D_r(t) F_+(t),$$

and the proof is completed. □

Obviously, Theorem 4.3 holds for the left indices as well.

It is easy also to check that all results of sections 3 and 4 remain true for the case of arbitrary bounded domain with rectifiable boundary instead of the unit disk.

5. Some remarks on the Birkhoff indices

As was mentioned in section 2, the Birkhoff indices are not determined uniquely. Here we consider the question how wide can be the set of different collections of the Birkhoff indices for a given matrix function $A(t) \in G\mathcal{R}^{n\times n}$. We begin with an estimation of the Birkhoff indices for matrix polynomial.

1^0. *If*

$$A(t) = \sum_{k=0}^{m} t^k A_k, \quad \det A(t) \neq 0 \quad (|t| = 1)$$

and $(b_j)_1^n$ *is a collection if Birkhoff indices of* $A(t)$ *then*

$$-(n-1)m \le b_j \le m. \tag{5.1}$$

Proof. Rewrite the equality (2.3) in the form

$$F_-^{-1}(t) t^{-m} A(t) = F_+(t) D_b(t) t^{-m}. \tag{5.2}$$

Suppose that $b_l > m$ for some l. An arbitrary element in the l-th column of the matrix function $F_+(t) D_b(t) t^{-m}$ has the form $f_{kl}^+(t) t^{b_l - m}$, hence it is analytic in $|t| \le 1$. The matrix function in the left-hand side of (5.2) belongs to $\mathcal{R}_-^{n\times n}$ and vanishes at ∞. Therefore $f_{kl}^+(t) t^{b_l - m} \equiv 0$ $(k = 1, 2, \ldots, n)$ which contradicts the invertibility of $F_+(t)$ $(|t| \le 1)$.

The left inequality (5.1) follows from the right one, because by (2.4)

$$b_j = \text{ind det } A(t) - \sum_{k\neq j} b_k \geq - \sum_{k\neq j} b_k \geq -(n-1)m.$$

□

Let us note that in the example (5.3) stated below the inequalities (5.1) are best possible (see 3^0).

2^0. *For any matrix function $A(t) \in G\mathcal{R}^{n\times n}$ the set of different collections of the Birkhoff indices is finite.*

For a matrix polynomial this assertion follows from 1^0. An arbitrary rational matrix function $A(t)$ can be written in the form

$$A(t) = q^{-1}(t)P(t)$$

where $P(t)$ is a matrix polynomial and $q(t)$ is a scalar polynomial. Hence every collection of Birkhoff indices of $A(t)$ has the form $(b_j - \text{ind } q(t))_1^n$ where $(b_j)_1^n$ is a collection of the Birkhoff indices of $P(t)$. □

Now we will show that the number of different collections of the Birkhoff indices for $A(t) \in G\mathcal{R}^{n\times n}$ can be arbitrary. We restrict ourselves to the simplest case $n = 2$.

Let

$$A(t) = \begin{bmatrix} 1 & 0 \\ p(t) & 1 \end{bmatrix} \tag{5.3}$$

where $p(t)$ is a polynomial vanishing in $t = 0$:

$$p(t) = \sum_{k=1}^{m} a_k t^k. \tag{5.4}$$

Obviously, the right indices and the left indices of $A(t)$ are $(0, 0)$. The pair $(0, 0)$ is also one of the pairs of its Birkhoff indices. If a pair (b_1, b_2) is another pair of the Birkhoff indices of $A(t)$, then $b_1 + b_2 = 0$, i.e., this pair is of the form $(n, -n)$ $(n \in \mathbf{Z})$. Represent $A(t)$ in the form

$$A(t) = \begin{bmatrix} t^{-n} & 0 \\ t^{-n}p(t) & t^n \end{bmatrix} \begin{bmatrix} t^n & 0 \\ 0 & t^{-n} \end{bmatrix} \tag{5.5}$$

It follows from (5.5) that the pair $(n, -n)$ is a pair of the Birkhoff indices of $A(t)$ if and only if the right indices of the matrix function

$$B(t) = \begin{bmatrix} t^{-n} & 0 \\ t^{-n}p(t) & t^n \end{bmatrix}$$

are zeros. It follows from [4, Theorem 1.2] that this is impossible for $n < 0$. On the other hand, it is proved in [4, Corollary 4.1] that for $n > 0$ the right indices of $B(t)$ are zero if and only if the Hankel matrix

$$H_n = \begin{bmatrix} a_1 & a_2 & a_3 & \cdots & a_n \\ a_2 & a_3 & a_4 & \cdots & a_{n+1} \\ a_3 & a_4 & a_5 & \cdots & a_{n+2} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ a_n & a_{n+1} & a_{n+2} & \cdots & a_{2n-1} \end{bmatrix}$$

is non-degenerate (here $\{a_j\}_1^m$ are the coefficients of the polynomial $p(t)$ and $a_j = 0$ for $j > m$). Now choose the numbers $\{a_j\}_1^m$ in the following way:

$$a_1 = a_2 = 1, \quad a_{n+1} = \frac{a_n^n}{n!} \quad (n = 2,\ 3, \dots,\ m-1). \tag{5.6}$$

Evidently, $\det H_1 \neq 0$, $\det H_2 \neq 0$ and for $n > 2$

$$|\det H_n| \geq a_n^n - \left|\sum_j s_j\right|,$$

where $\{s_j\}_1^{n!-1}$ are all terms of $\det H_n$ exept $(-1)^{\frac{n(n-1)}{2}} a_n^n$. Every $|s_j|$ is a product of n numbers a_j, and at least one of this factors is a_k for some $k > n$. Since

$$0 < a_{j+1} \leq a_j \leq 1 \quad (j = 1,\ 2, \dots,\ m-1),$$

then $|s_j| \leq a_k \leq a_{n+1}$, and for $n \leq m-1$

$$\det H_n \geq a_n^n - (n! - 1)a_{n+1} = a_{n+1} > 0.$$

Obviously, $|\det H_m| = a_m^m > 0$ and $\det H_n = 0$ for $n > m$.

Thus we obtain the following result.

3^0 *The set of all pairs of the Birkhoff indices for the matrix function* (5.3) *under the conditions* (5.4) *and* (5.6) *is following:*

$$(0,0),\ (1,-1), \dots,\ (m,-m).$$

Remark. If $p(t) = \sum\limits_{k=1}^{\infty} a_k t^k$, and all coefficients a_k $(k = 1,\ 2, \dots)$ are defined by (5.6) then we obtain non-rational (but entire) matrix function (5.3) which has infinite set of pairs of the Birkhoff indices $\{(k,-k)\ :\ k \geq 0\}$.

References

1. G. D. Birkhoff, *A theorem on matrices of analytic functions*, Math. Ann. **74** (1913), 122–133.

2. I. S. Čebotaru, *The reduction of systems of Wiener-Hopf equations to systems with vanishing indices*, Bull. Acad. Shtiince RSS Moldov. N 8 (1967), 54–66.(Russian)

3. K. Clancey and I. Gohberg, *Factorization of matrix functions and singular integral operators*, Birkhäuser, Basel, 1981.

4. I. Feldman, I. Gohberg and N. Krupnik, *On explicit factorization and application*, Integral Equations and Operator Theory **21** (1995), 430–459.

5. I. C. Gohberg and I. A. Feldman, *Convolution equations and projection methods for their solutions*, Amer. Math. Soc., Providence, R. I., 1974.

6. I. C. Gohberg and M. G. Krein, *Systems of integral equations on a half line with kernels depending on the difference of arguments*, Amer. Math. Soc. Transl. (2) **14** (1960), 217–287.

7. I. Gohberg, P. Lancaster and L. Rodman, *Invariant subspaces of matrices and applications*, Wiley, New-York, 1986.

8. R. Hagen, S. Roch and B. Silbermann, *Spectral theory of approximation methods for convolution equations*, Birkhäuser, Basel, 1995.

9. G. H. Hardy, J. E. Littlwood, G. Polya, *Inequalities*, Cambridge University Press, Cambridge, 1934.

10. G. S. Litvinchuk and I. M. Spitkovskii, *Factorisation of measurable matrix functions*, Birkhäuser, Basel, 1987.

11. A. W. Marshall and I. Olkin, *Inequalities: Theory of majorization and its applications*, Academic Press, New York, 1979.

I. Feldman
Dept of Math. and Comp. Sci.
Bar-Ilan University
52900 Ramat-Gan
Israel

A. Markus
Dept of Math. and Comp. Sci.
Ben-Gurion University of the Negev
84105 Beer-Sheva
Israel

MSC 1991: 47A68

Submitted: December 12, 1996