An Operator Identities Approach to Bezoutians. A General Scheme and Examples*

Vadim Olshevsky[†]and Lev Sakhnovich[‡]
Department of Mathematics
University of Connecticut
Storrs, CT 06269

Abstract

In this paper we propose a general scheme to study Bezoutians that is based on the method known as the *method of operator identities* in the operator theory literature [S76b, S97, S99] (its finite dimensional counterpart is known under the name *displacement structure method* in the engineering [KKM79, K99] and in matrix theory and numerical literature [HR84, O03]). The latter approach allows us to introduce a generalized concept of the *operator Bezoutian* and to carry over to it the classical results of Jacobi (on common roots of scalar polynomials [J1836]), and of Hermite (on polynomial stability [H1856]). Several other known results scattered in the mathematical and engineering literature (Krein [K] Krein [K], Sakhnovich [S76a], Gohberg-Heinig [GH76], Anderon-Jury [AJ76], Lerer-Tysmenetsky [LT82], Lerer-Rodman [LR96a, LR96b]) are shown to appear as particular instances of our general scheme. The unified *operator identities (displacement structure)* approach results in a transparent concise derivation of fairly general results allowing us to include, in one paper, most of known special cases (such as matrix polynomials and rational matrix functions) as well as some new (such as matrix analytic functions).

1 Introduction

1.1 Bezoutians. The Classical Results of Jacobi and Hermite

Bezoutian matrices were instrumental already in the classical studies of Euler [E1748] and Bezout [B1764] in the elimination theory, of Jacobi [J1836] and Sylvester [S1853] in the theory of separation of polynomial roots, and of Hermite [H1856] in polynomial stability. An excellent historical source is a survey paper [KN36][1], the references [H70], [W90] can also be consulted. The nomenclature "Bezoutian" is due to Sylvester [S1853], who introduced it using a definition that is no longer used nowadays[2]. The following, now standard, definition is due to Cayley .

Definition 1 ([C1857]) *Let*

$$F(z) = f_0 + f_1 z + f_2 z^2 + \ldots + f_n z^n, \qquad G(z) = g_0 + g_1 z + g_2 z^2 + \ldots + g_m z^m$$

---

[*] This work was supported in part by the NSF contracts 0242518 and 0098222.

[†] web page: http://www.math.uconn.edu/~olshevsky email: olshevsky@math.uconn.edu

[‡] email: vze3bh8z@verizon.net

[1] It is interesting to note the difference between the presentation of classical results on polynomial stability in [KN36] and [G53]. While [KN36] makes the Bezoutian matrix instrumental in its presentation, [G53] totally avoids the latter concept, thus presenting only a selection of the relevant results.

[2] Sylvester considers in [S1853] a special linear system of equations, and he associates with each of the n equations a function. He further writes: *"As these derived functions are of frequent occurrence, I find it necessary to give them a name, and I term them the n Bezoutics or Bezoutian primaries..."* He later adds: *"The determinant formed by arranging in a square the n sets of the coefficients of the n Bezoutians, and which I call the* **Bezoutian matrix***, gives, as is well known, the resultant."* (Along with the primary Bezoutians, he also considers functions associated with what we now call the U-factor [in the LU factorization] of the Bezoutian matrix calling them the *"secondary Bezoutians."*)

be two given polynomials, and assume for simplicity that $m \le n$. *Since the numerator in* $\rho(\lambda,\mu) = \frac{F(\lambda)G(\mu)-G(\lambda)gF(\mu)}{\lambda-\mu}$ *vanishes at* $\lambda=\mu$ *hence* $\rho(\lambda,\mu)$ *is a bivariate polynomial of degree* $n-1$. *Therefore, it can be written as*

$$\rho(\lambda,\mu) = \frac{F(\lambda)G(\mu)-G(\lambda)gF(\mu)}{\lambda-\mu} = \sum_{k=1}^{n-1} t_{i,j}\lambda^i\mu^j = \left[\begin{array}{cccccc} 1 & \lambda & \lambda^2 & \dots & \lambda^{n-1} \end{array}\right] T \left[\begin{array}{cccccc} 1 & \mu & \mu^2 & \dots & \mu^{n-1} \end{array}\right]^T. \tag{1.1}$$

The matrix $T = \left[\begin{array}{c} t_{i,j} \end{array}\right]$ *is called the Bezoutian of* $\{F(z), G(z)\}$.

The following two theorems are fundamental.

Theorem 2 (Jacobi-like) *Let* T *be the Bezoutian matrix, defined by* (1.1)*, of two scalar polynomials* $F(z)$ *and* $G(z)$. *Then*

$$\dim\text{Ker}T = \text{the number of common zeros of } F(z) \text{ and } G(z) \text{ counting with multiplicities.}$$

Theorem 3 (Hermite) *The polynomial*

$$P(x) = p_0 + p_1x + p_2x^2 + \cdots + p_nx^n \tag{1.2}$$

has all its roots in the open upper-half-plane if and only if the matrix $T = \left[\begin{array}{c} t_{k,l} \end{array}\right]$ *is positive definite, where the entries of the latter matrix are obtained from the expression*

$$-\frac{i}{2}\cdot\frac{P(\lambda)\breve{P}(\mu)-\breve{P}(\lambda)P(\mu)}{\lambda-\mu} = \sum_{k,l=0}^{n-1} t_{k,l}\lambda^k\mu^l = \left[\begin{array}{cccccc} 1 & \lambda & \lambda^2 & \dots & \lambda^{n-1} \end{array}\right] T \left[\begin{array}{cccccc} 1 & \mu & \mu^2 & \dots & \mu^{n-1} \end{array}\right]^T. \tag{1.3}$$

Here $\breve{P}(x) = {p_0}^* + {p_1}^*x + {p_2}^*x^2 + \cdots + {p_n}^*x^n$.

Remark 4 *Substituting the "split" polynomial* $P(z) = F(z) + iG(z)$ *back into (1.3) one immediately sees that the matrix* T *in (1.3) is exactly the Bezoutian matrix* T *of these* $F(z)$ *and* $G(z)$.

Theorem 3 first appeared in [H1856]. There are several reasons why it is more difficult to give only one reference for the theorem 2 as stated above. Namely, **(i)** it was not up until [C1857] that the definition 1 was given, and using another definition means a close but formally different statement; **(ii)** many authors worked not with Bezoutians but with the equivalent resultant matrices; **(iii)** often only the condition $\det T = 0$ was discussed, and the more "detailed" number $\dim\text{Ker}T$ was not needed; **(iv)** finally, most of the early authors used additional assumptions, e.g., that $F(z)$ and $G(z)$ have only simple roots. Still, one of the most relevant references in this regard [J1836] where Jacobi established that T is symmetric and

$$\left[\begin{array}{cccccc} 1 & \lambda & \lambda^2 & \dots & \lambda^{n-1} \end{array}\right] T \left[\begin{array}{cccccc} 1 & \lambda & \lambda^2 & \dots & \lambda^{n-1} \end{array}\right]^T = F(\lambda)G'(\lambda) - G(\lambda)F'(\lambda). \tag{1.4}$$

Jacobi used (1.4) to notice that $\det T = 0$ implies that $F(z)$ and $G(z)$ have common zeros, and the argument can be immediately adapted to yield theorem 2 at least when one of the polynomials has simple roots.

Householder points out that a proof of theorem 2 can be found in a later paper [D1876].

1.2 Further Generalizations

The work on the extension of these classical results progressed in three main directions.

1. Extensions to scalar functions. Perhaps the first reference where the concept of Bezoutian was introduced for something more than a scalar polynomial was [K36], where it was introduced for entire functions given in the series form. In [K] and in [GH76] it was used for the pair of entire functions of the form

$$F(z) = 1 + \int_0^w e^{izt}\overline{\Phi(t)}dt, \qquad \Phi(t) \in L_1(0,w).$$

In [S76a] the Bezoutian was introduced for the pair of entire functions of the form

$$F(z) = 1 + iz\int_0^w e^{izt}\overline{\Phi(t)}dt, \qquad \Phi(t) \in L_1(0,w).$$

In all these papers the authors obtained the analogs of the theorems 2, 3 for their classes of functions.

2. Extensions to matrix functions. Giving a definition of a Bezoutian for matrix polynomials is not immediate, if one wants to avoid Kronecker products and to deal with a moderately sized generalization. Anderson and Jury [AJ76] gave such a definition, and conjectured that the counterparts of the theorems 2 and 3 can be carried over their Bezoutian. This nontrivial task was carried in [LT82] by means of the theory just developed in [GLR82].

The Bezoutians for rational matrix functions were introduced in [LR96a, LR96b]. Lerer and Rodman proved the analogs of the theorems 2 and 3.

3. General approaches. An attempt go develop a general theory of Bezoutians was done in [1]. Haimovici and Lerer proposed a fairly general definition and showed that the Bezoutians of [AJ76, GH76, S76a] appear as special cases of their general scheme. They established a number of basic properties of the generalized Bezoutian; however, the counterparts of the theorems 2, 3 were not obtained.

1.3 Main Results

In this paper we continue the work of our colleagues. We introduce a generalized concept of the *operator Bezoutian* and carry over to it theorems 2, 3. We then apply the proposed general approach to all the above mentioned scalar and matrix special cases. Specifically, the results of Krein [K] Krein [K], Sakhnovich [S76a], Gohberg-Heinig [GH76], Anderon-Jury [AJ76], Lerer-Tysmenetsky [LT82], Lerer-Rodman [LR96a, LR96b] are shown to appear as particular instances of our general scheme.

We include more new examples, e.g., matrix analytic functions.

The unified approach is based on the method known as the *method of operator identities* in the operator theory literature [S76b, S97, S99] (its finite dimensional counterpart is known under the name *displacement structure method* in the engineering [KKM79, K99] and in matrix theory and numerical literature [HR84, O03]). It results in a transparent concise derivation of fairly general results allowing us to include, in one paper, many known special cases as well as some new.

2 A General Scheme

2.1 Definition and the Kernel Structure of the Operator Bezoutian T

Let H and G be Hilbert spaces; H being possibly infinite-dimensional, whereas G is finite dimensional: $m = \dim G$. In this paper we consider a pair of operator functions given in the form called *a realization*,

$$F(z) = I_m - zQ^*(I - Az)^{-1}\Phi, \tag{2.5}$$

$$G(z) = I_m - zP^*(I - Az)^{-1}\Phi, \tag{2.6}$$

where $A \in L(H); \Phi, P, Q \in L(G, H)$. Here $L(G, H)$ denotes the set of all linear bounded operators acting from G into H, and $L(H)$ stands for the set of all linear bounded operators acting from H into H. There are several types of realization formulas (cf. with [S99] and [BGR90]), we consider here the form shown in (2.5) since it is general enough to capture several special cases of our primary interest, two of them are mentioned next.

Example 1. Matrix entire functions of the exponential type. Let us specify A to be defined by

$$Af = i\int_0^x f(t)dt, \qquad f \in L_m^2(0, a) =: H.$$

It can be shown that in this case $F(z)$ and $G(z)$ are matrix entire functions of the exponential type.

Example 2. Matrix polynomials. If not only G but also H is finite dimensional as well, and if A has the Jordan canonical form then $F(z)$ and $G(z)$ are matrix polynomials.

Realization formulas are intrinsically connected with the method of operator identities [S96]. Using this method we obtain here the following results.

- A concept of common zeros of $F(z)$ and $G(z)$ for the matrix case $m > 1$ is introduced (It is trivial for the scalar case $m = 1$, but for the matrix case it is not).

- Our technique is based on introducing and analyzing the operator Bezoutian for the matrix entire functions $F(z)$ and $G(z)$.

  Note that Bezoutians of scalar polynomials are classical going back to the work of Hermite and Darboux and their connection to the root localization problems is well-understood (see, e.g., [KN36] and the references therein). Later this concept was extended to matrix polynomials [AJ76], [LT82], and to some classes of scalar entire functions [S76a, S96]. The definition of operator Bezoutian given here applies to the entire operator functions of the form (2.5) and (2.6).

In this section we introduce the operator Bezoutian T of matrix entire functions $F(z), G(z)$. To this end let us associate with the pair $F(z), G(z)$ in (2.5) , (2.6) the operator identity

$$TB - C^*T = N_2N_1^*, \qquad (T, B, C \in L(H); \quad N_1, N_2 \in L(G, H)) \tag{2.7}$$

where

$$B = A + \Phi Q^*, \qquad C = A + \Phi P^*, \tag{2.8}$$

and

$$N_1 = T^*\Phi, \tag{2.9}$$

We are ready to formulate the first result.

Theorem 5 *Assume that*

(i) *The relations (2.7), (2.8), and (2.9) hold true.*

(ii) *The two conditions (i) $M \in \mathrm{Inv}(A^*)$ and (ii) $M \subset \mathrm{Ker}\Phi^*$ imply $M = 0$ (i.e., there is no non-trivial A^*-invariant subspace annihilated by Φ^*).*

Then

$$\mathrm{Ker}T = L_1, \qquad \textit{where } L_1 \textit{ is the maximal } B\textit{-invariant subspace in } \mathrm{Ker}N_1^*. \tag{2.10}$$

2.2 Common zeros of the matrix entire functions $F(z)$ and $G(z)$. A generalization of the Jacobi result

Let us use (2.6) introduce

$$G_1(\lambda) = G(\frac{1}{\lambda}) = I_m + P^*(A - \lambda I)^{-1}\Phi. \tag{2.11}$$

Let λ_j be an eigenvalue if the operator C and let g_{q_j} be the corresponding generalized eigenvector of the order q_j.

Proposition 6 *If the operators A and C do not have common eigenvalues then*

$$P^*g_0 \neq 0. \tag{2.12}$$

Theorem 7 *Let the assumptions of theorem 5 and the proposition 6 are fulfilled, and let $m = 1$, i.e. $F(z)$ and $G(z)$ are scalar functions. If*

$$Q - P = N_1\Sigma, \qquad \Sigma \in L(G), \tag{2.13}$$

then the number of common zeros of the functions $F(z)$ and $G(z)$ coincides with $\dim \mathrm{Ker}T$.

Let us now consider the general matrix case $m > 1$, in this case one needs to explain what is meant by the number of common zeros. The standard definition is given next.

Definition 8 *Let all the generalized eigenvectors of the operator B have the order of 1. The matrix functions $F(z)$ and $G(z)$ are said to have the common zero $z_j = \frac{1}{\lambda_j}$ of the multiplicity κ_j if the following conditions are met:*

1. *There exist a $m \times \kappa_j$ matrix U_j of the full rank such that*

$$F(z_j)U_j = G(z_j)U_j = 0. \tag{2.14}$$

2. *The number κ_j is the maximal number for which a full-rank matrix U_j satisfying (2.14) can be found.*

With this definition one is able to extend the theorem 7 to the matrix case, we formulate first the simplest result when there are only eigenvectors and no generalized eigenvectors of higher oreders.

Theorem 9 *Let the conditions of the theorem 7 be fulfilled, and let all the generalized eigenvectors if the operator B have the order equal to 1. If*

$$Q - P = M_1\Sigma, \qquad \text{where} \qquad \Sigma \in L(G), \quad \det\Sigma \neq 0, \tag{2.15}$$

then the number of common zeros of the matrix functions $F(z)$ and $G(z)$ is equal to $dimKerT$.

Let us now treat the general case.

Definition 10 *We say that the matrix functions $F(z)$ and $G(z)$ have $\dim L$ common zeros, where L is the maximal subspace that is simultaneously B-invariant and C-invariant.*

It follows from the proof of Theorem 9 that in the case all the generalized eigenvectors of B have the order 1 the definitions 8 and 10 coincide.

Theorem 11 **(A generalization of the Jacobi result)** *Let conditions of the theorem 7 be fulfilled. If the relation (2.15) is valid, and*

$$\Phi g \neq 0, \qquad (g \neq 0), \tag{2.16}$$

then the number of common zeros of the matrix functions $F(z)$ and $G(z)$ is equal to $\dim\mathrm{Ker}T$.

2.3 Stability. A Generalization of the Hermite result

Let us consider only one matrix function

$$F(z) = I - zQ^*(I - Az)^{1}\Phi. \tag{2.17}$$

By setting

$$C = B, \qquad \text{and} \qquad N_2 = iN_1\Sigma, \qquad \text{where} \qquad \Sigma \in L(G), \qquad \Sigma \geq 0 \tag{2.18}$$

the relations (2.7) takes the form

$$TB - B^*T = iN_1\Sigma N_1, \qquad T = T^*, \qquad N_1 = T\Phi. \tag{2.19}$$

We assume here that $\lambda = 0$ is not a point of the limit spectrum of the operator T, implying

$$\dim\mathrm{Ker}T < \infty. \tag{2.20}$$

It follows from (2.19) that $\dim\mathrm{KerT}$ is simultaneously B-invariant and T-invariant. Hence H_1 is T-invariant and B^*-invariant. The operator T is bounded on H_1 together with its inverse. The spectrum of the operator

$$B_1 = P_1BP_1 \tag{2.21}$$

is contained in the spectrum of the operator B. Let us consider on H_1 the indefinite scalar product

$$[\varphi, \psi] = (T\varphi, \psi), \qquad \varphi, \psi \in H_1. \tag{2.22}$$

It follows from (2.19) that

$$[B_1\varphi, \psi] - [\varphi, B_1\psi] = i(\Sigma N_1^*\varphi, N_1^*\psi). \tag{2.23}$$

Thus, the operator B_1 is T-dissipative [K78]. Let us denote by $d_1(\lambda_j)$ the dimension of the root space of the operator B_1 with respect to the eigenvalue λ_j. We have [K78]

$$\sum_{Im\lambda_j<0} d_1(\lambda_j) \leq \kappa, \qquad \lambda_J = \frac{1}{z_j}, \tag{2.24}$$

where κ is the dimension of the maximal T-invariant subspace on which T is nonnegative.

Corollary 12 *If the operator A does not have the points of spectrum in the upper half plane ($Imz > 0$) and if the operator B_1 does not have real eigenvalues, then*

$$\sum_{Im\lambda_j<0} d_1(\lambda_j) = \kappa, \qquad \lambda_J = \frac{1}{z_j}, \tag{2.25}$$

i.e., the equality has place in (2.24).

Theorem 13 (A generalization of the Hermite, Schur-Cohn and Krein theorems) *Let the relations (2.19) be fulfilled and let*

$$T \geq \delta I, \qquad \delta 0). \tag{2.26}$$

Then $B = B_1$ and

$$\det F(z) \neq 0, \qquad Imz > 0. \tag{2.27}$$

References

[AJ76] B.D.O Anderson and E.I.Jury, *Generalized Bezoutian and Sylvester matrices in multivariable linear control*, IEEE Transactions on Automatic Control, AC-21 (1976), 551-556.

[K36] Krein in Akhiezer, Krein *Some questions in the theory of moments,*

[B1764] $\acute{E}$.Bezout, *Recherches sur le degré des équations résultantes de l'evanouuissement des inconnuues, et sur les moyens qu'il convenient d'employer pour trouver ces équations*, Mém Acad. Roy. Sci Paris, 1764, 288-338.

[BGR90] J. Ball, I.Gohberg and L.Rodman, *Interpolation of rational matrix functions*, OT45, Birkhäuser Verlag, Basel, 1990.

[C1857] A.Cayley, *Note sur la méthode d'élimination de Bezout*, J. Reine Angew. Math., **53** (1857), 366-367.

[D1876] Gaston Darboux, *Sur la téorie l'élimination entre deux équations à une inconnue*, Bull Sci. Math., 10 (1876), 56-64.

[E1748] L.Euler, *Introductio in Analysin Infinitorum*, Tom 2, Lausanne, 1748.

[G53] F.R. Gantmacher, The Theory of Matrices, Chelsea, New York, 1974.
First Russian edition was published in Moscow, 1953 by Gosudarstv. Izdat. Tehn.-Teor. Lit.

[GH76] I.Gohberg and G.Heinig, *The Resultant Matrix and its Generalizations. II. Continual Analogs of Resultant Matrix*, Acta Math. Acad. Sci. Hungar. , **28**, (1976), 198-209.

[GLR82] I.Gohberg, P.Lancaster and L.Rodman, *Matrix Polynomials*, Academi Press, 1982.

[H70] A.S.Householder, *Bezoutiants, elimination and localization*, SIAM Rev. 12 (1970) 73–78.

[H1856] Hermite, C., *Extrait d'une lettre de Mr. Ch. Hermite de Paris à Mr. Borchardt de Berlin, sur le nombre des racines d'une èquation algèbrique comprises entre des limits donèes,* J. Reine Angew. Math., **52** (1856), 39-51.

[1] [HL95] I.Haimovic and L.Lerer, *Bezout Operators for Analytic Operator Functions, I. A General Concept of Bezout Operator,* Integral Equations and Operator Theory, **2** (1995), 33-70.

[HR84] G. Heinig and K. Rost, *Algebraic methods for Toeplitz-like matrices and operators*, Berliin, Academie-Verlag (1984).

[J1836] G.G.J.Jacobi, *De eliminatione variabilis e duabus aequationibus algebraicis*, J. Reine Angew Math., **15** (1836), 101-124.

[K] M.G.Krein, Unpublished Manuscript. The Krein result was cited in [S76a].

[K78] M.G.Krein, *Introduction to the Geometry of Indefinite J-spaces and to the Theory of Operators in those Spaces*, American Mathematical Society Translations (2), vol. 93, 1978, ???-176.

[K99] T. Kailath, *Displacement structure and array algorithms*, In: *Fast reliable algorithms for matrices with structure*, (T. Kailath, A.H. Sayed, Eds.),SIAM Publications, Philadelphia 1999.

[KKM79] T.Kailath, S.Kung and M.Morf, *Displacement ranks of matrices and linear equations*, J. Math. Anal. and Appl., **68** (1979), 395-407.

[KN36] M.G.Krein and M.A.Naimark, *The Method of Symmetric and Hermitian Forms in the Theory of Separation of the Roots of Algebraic Equations*, Linear and Multilinear Algebra, 1981, **10**, 265-308. (originally published in Russian in Kharkov, 1936).

[LT82] L.Lerer and M.Tismenetsky,*The Bezoutian and Eigenvalue-Separation Problem for Matrix Polynomials*, Integral Equations and Operator Theory, **5**, (1982), 386-445.

[LR96a] L.Lerer and L.Rodman, *Bezoutians of Rational Matrix Functions,* Journal of Functional Analysis, Vol 141, No.1 (1996), 1-36.

[LR96b] L.Lerer and L.Rodman, *Symmetric factorizations and localization of zeroes of rational matrix functions*, Linear and Multilinear Algebra 40 no. 3 (1996), 259-281.

[O03] V. Olshevsky, *Pivoting for structured matrices and rational tangential interpolation*, in *Fast Algorithms for Structured Matrices: Theory and Applications* (V.Olshevsky, editor), CONM/323, p. 1 - 75, AMS publications, May 2003.

[S1853] I.Sylvester, *On a Theory of Syzygetic Relations of two Rational Integral Functions, Comprising an Application to the Theory of of Sturm's Functions, and That of the Greatest Algebraical Common Measure*, Phylos. Trans. Roy. Soc. London, **143** (1853), 407-548.

[S76a] L.A.Sakhnovich, *The Operator Bezoutian in the Theory of the Roots of Entire Functions,* Functional Analysis Appl., **10**, 45-51 (1976). In Russian.

[S76b] L.A.Sakhnovich, *On the Factorizatrion of an Operator-valued Transfer Function,* Sov. Math. Dokl., **17**, 203-207 (1976). In Russian.

[S86] L.Sakhnovich, *Factorization problems and operator identities*, Russian Mathematical Surveys, **41 : 1** (1986), 1 - 64.

[S96] L.A.Sakhnovich,*Integral Equations with Difference Kernels on Finite Intervals,* Operator Theory Series, v.84, Birkhauser Verlag, 1996.

[S97] L.A.Sakhnovich *Interpolation Theory and its Applications,* Kluwer Academic Publications, 1997.

[S99] L.A.Sakhnovich, *Spectral Theory of Canonical Differential Systems, Method of Operator Identities,* Operator Theory Series, v. 107, Birkhauser Verlag, 1999.

[W90] H.Wimmer, *On the history of the Bezoutian and the resultant matrix*, Linear Algebra and its Applications, **128**, 1990, 27-34.