Letters in Mathematical Physics **39:** 349–353, 1997.

Matrix Vieta Theorem Revisited

ALAIN CONNES[1] and ALBERT SCHWARZ[2*]
[1]*College de France, IHES, Bures-sur-Yvette, F91440, France. e-mail: connes@ihes.fr*
[2]*Department of Mathematics, University of California, Davis, CA 95616, U.S.A.*
e-mail: schwarz@math.ucdavis.edu.

(Received: 26 February 1996)

Abstract. We give another proof of the noncommutative analog of the Vieta theorem. This proof gives a little bit stronger statement and leads to some generalizations.

Mathematics Subject Classifications (1991): 15A24, 16B.

Key words: Vieta theorem, matrix equations, symmetric functions.

The following matrix analog of the Vieta theorem was proved in [1]. Let $X_1, ..., X_n$ denote n independent solutions of the matrix equation

$$X^n + A_1 X^{n-1} + \cdots + A_n = 0. \tag{1}$$

Here $X, A_1, ..., A_n$ are $m \times m$ complex matrices. (The solutions $X_1, ..., X_n$ are called independent if the coefficients $A_1, ..., A_n$ can be expressed in terms of $X_1, ..., X_n$. A more precise definition of independence will be given later.) Then

$$\mathrm{Tr} X_1 + \cdots + \mathrm{Tr} X_n = -\mathrm{Tr} A_1, \tag{2}$$

$$\det X_1 \cdot \cdots \cdot \det X_n = (-1)^n \det A_n. \tag{3}$$

More generally,

$$\det(1 - \lambda X_1) ... \det(1 - \lambda X_n) = \det(1 + A_1 \lambda + \cdots + A_n \lambda^n) \tag{4}$$

for every complex λ. One can expand Equation (4) in power series with respect to λ. Then we obtain expressions for $\sum \mathrm{Tr} X_i^k$ in terms of the coefficients $A_1, ..., A_n$:

$$\sum \mathrm{Tr} X_i^k = \sum_{i_1 + \cdots + i_s = k} \frac{k}{s} \mathrm{Tr}(A_{i_1} ... A_{i_s})(-1)^s. \tag{5}$$

It was shown in [1] that analogous statements are correct also in the case when $X, A_1, ..., A_n$ in Equation (1) are considered as elements of an arbitrary associative ring $\mathcal{A}$ with trace and /or determinant satisfying the standard requirements. The

* Research supported in part by NSF grant No. DMS-9500704.

aim of this Letter is to give a new proof of the results of [1] and to present some generalizations of these results. Equations (2)–(4) were proved in [1] at first for the matrix case; the case of an arbitrary ring was analyzed by means of reduction to the matrix case. We were able to directly analyze the general case. Moreover, already in the matrix case, our statements are a little bit stronger than the results of [1].

Let us consider Equation (1), where $X, A_1, ..., A_n$ are elements of an associative ring $\mathcal{A}$. We say that the solutions $X_1, ..., X_n$ are independent if the coefficients $A_1, ..., A_n$ can be expressed in terms of $X_1, ..., X_n$. In other words, the matrix

$$\Xi = \begin{pmatrix} X_1^{n-1} & ... & X_n^{n-1} \\ \\ X_1 & ... & X_n \\ 1 & ... & 1 \end{pmatrix} \tag{6}$$

should be invertible (i.e. $\Xi \in \mathrm{GL}_n(\mathcal{A})$). We will prove that in this case the matrix

$$\alpha = \begin{pmatrix} -A_1 & ... & -A_n \\ 1 & & \\ & \ddots & \\ & & 1 \end{pmatrix} \tag{7}$$

is similar to the diagonal matrix

$$\xi = \begin{pmatrix} X_1 & & \\ & \ddots & \\ & & X_n \end{pmatrix}. \tag{8}$$

The proof is based on the remark that $\alpha\Xi = \Omega$, where

$$\Omega = \begin{pmatrix} X_1^n & .... & X_n^n \\ & .... & \\ X_1^2 & .... & X_n^2 \\ X_1 & .... & X_n \end{pmatrix}.$$

At the same time, we have $\Xi\xi = \Omega$, therefore

$$\alpha\Xi = \Xi\xi \tag{9}$$

(Here α, ξ, Ξ are considered as elements of the ring $M_n(\mathcal{A})$ of $n \times n$ matrices with entries from $\mathcal{A}$.) We assumed that Ξ is invertible, therefore α and ξ are similar.

Notice, that the identity (9) is similar to the key step of the proof of the Bott periodicity theorem (see, e.g., [5]).

Let us assume that there is a notion of determinant in the ring $\mathcal{A}$ (in other words, we fix a homomorphism det : $\mathcal{A} \to k$ where k is a commutative ring). It is well known [3] that in this case there is also a notion of determinant in the ring $M_n(\mathcal{A})$.

We obtain that

$$\det \alpha = \det \xi. \tag{10}$$

Equation (3) follows immediately from (10).

Now let us assume that there is a notion of a trace in the ring $\mathcal{A}$ (i.e. we fix an additive homomorphism Tr of $\mathcal{A}$ into an Abelian group obeying Tr $ab =$ Tr ba). Then we can introduce the notion of trace also in the ring $M_n(\mathcal{A})$. We obtain that

$$\mathrm{Tr}\, \alpha^l = \mathrm{Tr}\, \xi^l \tag{11}$$

for every $l = 1, 2, ...$. Equation (5) follows immediately from (11).

The statements above can be expressed in terms of noncommutative generalization of elementary symmetric functions. Let us consider the expressions of the coefficients $A_1, ...A_n$ of (1) in terms of independent solutions $X_1, ..., X_n$. These expressions (or, more precisely, the expressions for $\sigma_k = (-1)^k A_k$) are called elementary symmetric functions of noncommuting variables $X_1, ..., X_n$ (cf. [2]). Then it follows from the consideration above that

$$\sigma_n(X_1, ..., X_n) = X_1...X_n \quad \text{in } K_1(\mathcal{A}). \tag{12}$$

(Recall that $K_1(\mathcal{A})$ is a commutative group obtained from $\mathrm{GL}_\infty(\mathcal{A})$ by means of factorization with respect to the ideal generated by all elements of the form $aba^{-1}b^{-1}$.) Analogously,

$$\sigma_1(X_1, ..., X_n) = X_1 + \cdots + X_n \quad \text{in } \mathrm{HH}_0(\mathcal{A}), \tag{13}$$

where $\mathrm{HH}_0(\mathcal{A})$ can be defind as a quotient of $\mathcal{A}$ with respect to a linear subspace spanned by all elements of the form $ab - ba$, where $a, b \in \mathcal{A}$. (The group $\mathrm{HH}_0(\mathcal{A})$ can be identified with the zero-dimensional Hochschild homology group of $\mathcal{A}$.)

Recently, Gelfand and Retakh [4], motivated by a wish to find more complete formulation of the noncommutative Vieta theorem, than in [1], proved the following interesting theorem:

THEOREM. *For generic $X_1, ...X_n$ one can find rational expressions $v_l(X_1, ...X_n)$ in such a way that elementary symmetric functions $\sigma_k(X_1, ...X_n)$ can be expressed in terms of $Y_l = v_l^{-1} X_l v_l$ by the formula*

$$\sigma_k(X_1, ..., X_n) = \sum_{i_1 < i_2 < ... < i_k} Y_{i_k} Y_{i_{k-1}} ... Y_{i_1}.$$

It is easy to see that (12) and (13) follow from this theorem.

The above consideration can be generalized as follows. Let us consider a matrix

$$\alpha = \begin{pmatrix} A_{11} & ... & A_{1n} \\ & & \\ A_{n1} & ... & A_{nn} \end{pmatrix} \in M_n(\mathcal{A}). \tag{14}$$

An element $\Lambda \in \mathcal{A}$ is called an eigenvalue of α if there exist elements $x_1, ..., x_n \in \mathcal{A}$ obeying

$$\begin{aligned} A_{11}x_1 + \cdots + A_{1n}x_n &= x_1\Lambda \\ \cdots \qquad & \quad \cdots \\ A_{n1}x_1 + \cdots + A_{nn}x_n &= x_n\Lambda. \end{aligned} \tag{15}$$

The column

$$X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \tag{16}$$

is called an eigenvector of α.

We say that the eigenvectors

$$X_1 = \begin{pmatrix} x_{11} \\ \vdots \\ x_{n1} \end{pmatrix}, ..., X_n = \begin{pmatrix} x_{1n} \\ \vdots \\ x_{nn} \end{pmatrix} \tag{17}$$

are independent if the matrix

$$\Xi = (X_1, ..., X_n) = \begin{pmatrix} x_{11} & ... & x_{1n} \\ & & \\ x_{n1} & ... & x_{nn} \end{pmatrix} \tag{18}$$

is invertible.

One can verify, that every solution X of Equation (1) generates an eigenvector $(X^{n-1}, X^{n-2}, ..., 1)$ of the matrix (7); the corresponding eigenvalue is equal to X. If the solutions $X_1, ...X_n$ of equation (1) are independent, then the corresponding eigenvectors of (7) are also independent.

It is easy to check that

$$\alpha\Xi = \Xi\Lambda \tag{19}$$

where Λ stands for the diagonal matrix $\Lambda = \text{diag}(\Lambda_1, ..., \Lambda_n)$; α and Ξ are correspondingly given by (14) and (18). If the eigenvalues $\Lambda_1, ..., \Lambda_n$ correspond to independent eigenvectors, we obtain that the matrix α is similar to the matrix $\Lambda = \text{diag}(\Lambda_1, ..., \Lambda_n)$. This statement permits us to generalize the above results. For example, if there is a notion of trace in $\mathcal{A}$, we obtain that

$$\text{Tr } A_{11} + \cdots + \text{Tr } A_{nn} = \text{Tr } \Lambda_1 + \cdots + \text{Tr}\Lambda_n. \tag{20}$$

More generally,

$$\sum_{k=1}^{n} \mathrm{Tr}\, \Lambda_k^l = \sum_{1\leqslant i_r \leqslant n} \mathrm{Tr}\, A_{i_0 i_1} A_{i_1 i_2} ... A_{i_{l-1} i_0} \tag{21}$$

(We use the fact that the matrix α^l is similar to the diagonal matrix $\Lambda^l =$ diag $(\Lambda_1^l, ..., \Lambda_n^l)$. The equation (5) follows immediately from (21)).

Let us notice in conclusion that the matrix equation (1) is related to the matrix differential equation

$$\frac{\mathrm{d}^n \varphi}{\mathrm{d}t^n} + A_1 \frac{\mathrm{d}^{n-1}\varphi}{\mathrm{d}t^{n-1}} + \cdots + A_{n-1}\frac{\mathrm{d}\varphi}{\mathrm{d}t} + A_n \varphi = 0 \tag{22}$$

Namely, if $X_1, \ldots, X_n$ are independent solutions of (1), than the general solution of (22) can be written in the form

$$\varphi(t) = \mathrm{e}^{X_1 t} C_1 + \cdots + \mathrm{e}^{X_n t} C_n,$$

where $C_1, ..., C_n$ are arbitrary matrices. (The condition of independence permits us to express $C_1, ..., C_n$ in terms of Cauchy data.) Analogously, the system (15), where A_{ij} and x_i are considered as matrices, is related to the system of matrix differential equations

$$\begin{aligned} A_{11}\tfrac{\mathrm{d}\varphi_1}{\mathrm{d}t} + \cdots + A_{1n}\tfrac{\mathrm{d}\varphi_n}{\mathrm{d}t} &= 0, \\ A_{n1}\tfrac{\mathrm{d}\varphi_1}{\mathrm{d}t} + \cdots + A_{nn}\tfrac{\mathrm{d}\varphi_n}{\mathrm{d}t} &= 0. \end{aligned} \tag{23}$$

Acknowledgement

We are indebted to D. Fuchs, I. Gelfand and J. Loday for useful discussions.

References

1. Fuchs, D. and Schwarz, A.: *Matrix Vieta Theorem*, Amer. Math. Soc. Transl. ser. 2, vol. 169, Amer. Math. Soc., Providence, 1995.
2. Gelfand, I.M., Krob, D., Lascous, A., Leclerc, B., Retakh, V.S. and Thibon, J.-Y.: Non-commutative symmetric functions, Preprint, 1994, hep-th/9407124.
3. Milnor, J.: *Introduction to Algebraic K-Theory*, Ann. of Math. Stud. Princeton, Univ. Press, Princeton, 1971.
4. Gelfand, I.M. and Retakh, V.S.: Non-commutative Vieta theorem and symmetric functions, Preprint, 1995.
5. Atiyah, M.: *K-theory*, Harvard University Press, Cambridge, Mass., 1965.