Počátky teorie matic v Českých zemích a jejich ohlasy

Summary

In: Martina Štěpánová (author): Počátky teorie matic v Českých zemích a jejich ohlasy. (English). Praha: Matfyzpress, vydavatelství Matematicko-fyzikální fakulty UK, Praha, 2014. pp. 391–409.

Persistent URL: http://dml.cz/dmlcz/403399

Summary

Origins of Matrix Theory in Czech Lands (and the responses to them)

Matrix theory is a relatively new branch of mathematics. Its origins can be placed in the second half of the 19th century; some of the now familiar concepts were not conceived until the 20th century. The instigations leading to the birth of matrix theory were coming from several directions and their roots are centuries old. The notion matrix can be found about 2000 years ago in China in connection with the solution of systems of linear equations. After this came a long (approximately 1800 years) period when matrices were not much studied. Matrix theory did not develop directly from the problem of solving systems of linear equations but rather originated from the theory of determinants, bilinear and quadratic forms and the theory of linear substitutions.

It was the British mathematician James Joseph Sylvester in 1850, who established the modern conception of (rectangular) matrix, and used the term "matrix" for the first time in his article *Additions to the articles "On a new class of theorems", and "On Pascal's theorem"* [Sy2]. In 1858, the British mathematician Arthur Cayley, who was Sylvester's friend, published his article *A memoir on the theory of matrices* [Cy3] in the journal Philosophical Transactions of the Royal Society of London. We usually consider this article to be the origin of matrix theory. It is very striking that the beginning of the theory of determinants (1750) predated the origin of matrix theory.

Cayley's *Memoir* piqued little immediate interest in Great Britain and it was practically unknown elsewhere. Moreover, in the second half of the 19th century many mathematicians gave priority to determinants and it took a long time until matrix theory became an independent discipline. The process of gradual acceptance of this theory by mathematical community began in the second half of the 19th century and continued during the first half of the 20th century. Until then, a majority of mathematicians formulated their results, which are presently well-incorporated in matrix theory, in terms of determinants, bilinear and quadratic forms.

Textbooks published between 1860 and 1880 by Václav Šimerka, Josef Smolík, František Josef Studnička, and František Machovec belong among the first Czech publications on algebra.

Reseach in the fields which are presently classified as belonging to linear algebra corresponded in the Czech lands with the situation abroad – the oldest works on linear algebra which were published by Czech mathematicians are concerned with determinants, not with matrices. Their authors are František Josef Studnička, Karel Zahradník, Eduard Weyr, Matyáš Lerch, Wilhelm Matzka, etc. It is necessary to say that these were mostly shorter studies, which did not bring any significant or original ideas. Rather they revised results already known abroad and evolved minor improvements of the theory or its application in other branches of mathematics.

The first Czech mathematician using the matrices was most probably Ludvík Kraus. During his studies in Munich and Berlin he attended lectures of Felix Klein, Karl Theodor Wilhelm Weierstrass and Leopold Kronecker. Since the two latter named mathematicians belonged to the most noted personalities in this field, we can assume that the first basic concepts of matrix theory Ludvík Kraus acquired actually in Berlin. Kraus's early death, which was a great loss for the Czech mathematical community, prevented publication of any of his work relating to the subject of matrices.

It was probably Kraus who inspired an interest in matrices in the Prague mathematician Eduard Weyr, who later reached a world's recognition in this field.

Eduard Weyr was one of the few mathematicians on the European continent acquainted with matrix theory and working in this area at the time.

Eduard Weyr was born in Prague on June 22nd, 1852. His father František Weyr was a famous professor of mathematics at a Realschule in Prague. Eduard had nine siblings, his older brother Emil (1848–1894) was also an outstanding mathematician and professor at the University of Vienna.

Eduard Weyr studied at the Prague Polytechnic. He also attended lectures abroad. Among his teachers were Rudolf Friedrich Alfred Clebsch and Ernst Friedrich Wilhelm Klinkerfues in Göttingen, Charles Hermite and Joseph Alfred Serret in Paris, Leopold Kronecker and Immanuel Lazarus Fuchs in Berlin. In 1873, Eduard Weyr obtained a doctor's degree at Göttingen. In the academic year 1874/75, he habilitated at the Czech Polytechnic (research in geometry). The next year he was named a salaried docent and in 1876, he became an extraordinary professor at the Czech Polytechnic and also a private docent at the University of Prague. In 1891, he became a substitute professor at the Czech University in Prague. In the academic years 1884/85 and 1890/91, he held the post of a rector of the Czech Polytechnic. Eduard Weyr died in Záboří nad Labem on July 23rd, 1903.

During his life, Eduard Weyr was regarded mainly as a geometer, but was also concerned with analysis. Nowadays, his results in linear algebra, matrix theory and his study of the convergence of power series in matrices are more valued than his contributions to geometry.

Eduard Weyr published his original worldwide reputable results (see the section on the Weyr theory of characteristic numbers below) in matrix theory in the 1880s and early 1890s. He discovered the *Weyr characteristic*, which is a dual sequence to the better known Segre characteristic, and also the so-called *typical form*. This canonical form of a matrix is nowadays called the *Weyr canonical form* and it is permutationally similar to the commonly used Jordan canonical form of the same matrix. There are mathematical problems in which the solution is rather smarter using the Weyr canonical form and not the Jordan canonical form. In some instances the application of the Jordan canonical form would not lead to the solution because it does not have (contrary to the Weyr canonical form) the essential properties needed.

The following list is a chronological order of Weyr's works involving matrix calculus:

- *O základní větě v theorii matric* [We2], 1884,
- *Sur la théorie des matrices* [We5], 1885,
- *Répartition des matrices en espèces et formation de toutes les espèces* [We6], 1885,
- *O binarných matricích* [We8], 1887,
- *Sur la réalisation des systèmes associatifs de quantités complexes à l'aide des matrices* [We9], 1887,
- *O theorii forem bilinearných* [We12], 1889,
- *Zur Theorie der bilinearen Formen* [We13], 1890,
- *O theorii forem bilinearných* [We17], 1901.

Because of the connections between matrices and quaternions we can also mention the paper

- *Sur la théorie des quaternions* [We3], 1884.

There is no mention of the term "matrix" in the paper

- *Note sur la théorie des quantités complexes formées avec n unités principales* [We10], 1887,

which deals with linear associative algebras. Nevertheless, one may translate its results into terms of matrices. Eduard Weyr did it in his book *O theorii forem bilinearných* (1889).

The paper *O základní větě v theorii matric* is concerned with the proof of the Cayley-Hamilton theorem.

In the two 1885 papers *Sur la théorie des matrices* [We5] and *Répartition des matrices en espèces et formation de toutes les espèces* Weyr introduced the basic ideas of his very modern and original *theory of characteristic numbers* and discovered the so-called *typical form*. He described a complete system of invariants for matrix similarity, which consists of the set of all eigenvalues of a matrix, and the *Weyr characteristic of the matrix*. The Weyr characteristic of a matrix A contains all Weyr characteristics of A associated with all eigenvalues of A. The Weyr characteristic associated with λ is the dual sequence to the better known Segre characteristic associated with λ. The *typical form* is a canonical form of a matrix; nowadays, it is used in a slightly altered form and called the *Weyr canonical form*. In these short notes, Eduard Weyr also studied conditions for the existence of a minimal polynomial of degree smaller than the order of a given matrix, and conditions for the diagonalization of a matrix. He obtained the estimation for the nullity of the product of two matrices.

Weyr developed the basic ideas of his theory in the 111-page book *O theorii forem bilinearných* as well as in its German version *Zur Theorie der bilinearen Formen* in more detail.[1] The titles of the texts are confusing. These works are primarily concerned with matrix theory, while the theory of bilinear form is mentioned only on few pages. Weyr was one of the first mathematicians who contributed to connections between matrix theory and the theory of bilinear forms.

The German version *Zur Theorie der bilinearen Formen* was published in the journal *Monatshefte für Mathematik und Physik* and it is not a literal translation. Some parts are expanded and, on the other hand, some parts are abbreviated or missing.

The German paper *Zur Theorie der bilinearen Formen* and the French note *Répartition des matrices en espèces et formation de toutes les espèces* are the most cited Weyr's works in contemporary articles and monographs.

The 1887 paper *Sur la réalisation des systèmes associatifs de quantités complexes à l'aide des matrices* is focused on the connections between matrices and linear associative algebras.

In 1887, Eduard Weyr also published the monograph *O binarných matricích*, which was intended for the Czech mathematical community in order to support the new theory, introduce its notation and basic properties of matrices. Surprisingly, this work deals only with matrices of second order.

Weyr's works represented the only important contribution to matrix theory by Czech mathematicians in many decades that followed.

There are few Czech papers on matrix theory which were published during the first decades of the 20th century. Unfortunately, articles by Karel Petr, Václav Simandl, Václav Hruška, Jaroslav Jarušek, and Karel Rössler contained neither original nor significant results. The same is true of works which were written by Bohumil Bydžovský. Nevertheless, it is worth paying special attention to his book *Základy teorie determinantů a matic a jich užití* [By1] which was published in 1930.[2] It is one of the oldest monographs in the world whose title contains the word "matrix". This publication was important because of its educational function, Czech terminology and notation.

Direct and more comprehensive responses from Czech mathematicians to Weyr theory of matrices were written in Brno, mainly in the 1950s. The best known of the followers of Weyr in this Moravian town is Otakar Borůvka. At the end of World War II, he began to direct the research of Masaryk University towards differential equations. In the academic year 1946/47 he set up a Differential equations seminar. He and his colleague Jiří Čermák used Weyr theory to solve systems of linear differential and difference equations.

---

[1] The 1891 note *O theorii forem bilinearných* is a Weyr's short, 3-page contribution to a Prague congress.

[2] The second edition of this book appeared in 1947 under the title *Úvod do teorie determinantů a matic a jich užití*.

Borůvka also wrote the lecture notes *Matice* [Bo4] in 1947 (2nd edition: 1948, 3rd edition: 1966) and the textbook *Základy teorie matic* [Bo8] in 1971. The latter is the first Czech book explaining Weyr theory.

The approach of Miroslav Novotný, another Brno mathematician, was rather unusual for that time. He tried to present Weyr theory in a quite abstract language (in terms of the so-called *projective A-spaces* and their *A-collineations*).

Other more significant reactions to Weyr's outcomes appeared in our lands again after several decades. In the last decade of the 20th century a team of authors published the monograph *Eduard Weyr (1852–1903).* The book was initiated by Jindřich Bečvář who also contributed the majority of the text contained in the monograph. Besides the biographical part devoted to Eduard Weyr and his family, Jindřich Bečvář also compiled list of publications of this mathematician and above all he reformulated Weyr's about 100 year's old theory of characteristic numbers into the modern language of vector spaces and homomorphisms.

* * *

Weyr theory of characteristic numbers

The basic ideas of the *Weyr theory of characteristic numbers* as well as the relationship between the Weyr characteristic and the Segre characteristic are the subject of this section.

1 Remark. We shall always assume that A is a complex matrix. We denote the rank of a matrix A by $r(A)$.

2 Definition. Let A be a square matrix of order n. The set of all vectors v which satisfy $Av^T = o^T$ is called the *nullspace* or *kernel of A*. This set forms a subspace of $\mathbb{C}^n$ and we denote it by $\operatorname{Ker} A$.

3 Definition. Let A be a square matrix of order n. Then the positive integer

$$\operatorname{null} A = n - r(A)$$

is called the *nullity of A*.

The nullity of A is equal to the dimension of $\operatorname{Ker} A$.

4 Remark. Let A be a square matrix and λ its eigenvalue. Obviously, a matrix $A - \lambda E$ is singular. Hence, $\text{null}\,(A - \lambda E) > 0$. Moreover

$$\text{null}\,(A-\lambda E)^k \le \text{null}\,(A-\lambda E)^{k+1} \qquad \text{for all} \quad k = 0,\, 1,\, \ldots$$

If s is the multiplicity of λ, then $\text{null}\,(A-\lambda E)^k = \text{null}\,(A-\lambda E)^{k+1}$ if and only if $\text{null}\,(A-\lambda E)^k = \text{null}\,(A-\lambda E)^{k+1} = s$. Thus, there exists a positive integer $t \ge 1$ such that

$$0 < \text{null}\,(A-\lambda E) < \text{null}\,(A-\lambda E)^2 < \ldots$$
$$\ldots < \text{null}\,(A-\lambda E)^t = \text{null}\,(A-\lambda E)^{t+1} = \ldots,$$

and

$$s = \text{null}\,(A-\lambda E)^t = \text{null}\,(A-\lambda E)^{t+1} = \ldots$$

5 Definition. Let A be a square matrix and let λ be its eigenvalue. The smallest positive integer t satisfying

$$\text{null}\,(A-\lambda E)^t = \text{null}\,(A-\lambda E)^{t+1}$$

is called the *index of A associated with the eigenvalue λ.*

6 Definition. Let A be a square complex matrix, let λ be its eigenvalue and let t be the index of A associated with λ. The *characteristic numbers* of A associated with λ are defined as the positive integers

$$\begin{aligned}
&\eta_1 = \text{null}\,(A-\lambda E),\\
&\eta_2 = \text{null}\,(A-\lambda E)^2 - \text{null}\,(A-\lambda E),\\
&\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\\
&\eta_t = \text{null}\,(A-\lambda E)^t - \text{null}\,(A-\lambda E)^{t-1}.
\end{aligned}$$

The sequence $\eta(\lambda)$ of positive integers $\eta_1,\, \eta_2,\, \ldots,\, \eta_t$ is called the *Weyr characteristic of A associated with λ.* We write $\eta(\lambda) = (\eta_1,\, \eta_2,\, \ldots,\, \eta_t)$.

7 Theorem. *Let A be a square complex matrix, let λ be its eigenvalue of multiplicity s and let $\eta(\lambda) = (\eta_1,\, \eta_2,\, \ldots,\, \eta_t)$ be the Weyr characteristic of A associated with λ. Then*

(i) $\eta_1 \ge \eta_2 \ge \ldots \ge \eta_t > 0$,
(ii) $\eta_1 + \eta_2 + \ldots + \eta_t = s$.

8 Definition. Let $\lambda_1,\, \lambda_2,\, \ldots,\, \lambda_u$ be mutually distinct eigenvalues of a complex square matrix A, let $(\alpha_1,\ \alpha_2,\ \ldots,\ \alpha_{t_1})$ be the Weyr characteristic of A associated with the eigenvalue λ_1, let $(\beta_1,\ \beta_2,\ \ldots,\ \beta_{t_2})$ be the Weyr characteristic of A associated with the eigenvalue λ_2 etc. and let $(\upsilon_1,\ \upsilon_2,\ \ldots,\ \upsilon_{t_u})$ be the

Weyr characteristic of A associated with the eigenvalue λ_u. Then the system of positive integers

$$\eta(A) = [(\alpha_1,\ \alpha_2,\ \ldots,\ \alpha_{t_1}),\ (\beta_1,\ \beta_2,\ \ldots,\ \beta_{t_2}),\ \ \ldots,\ (\upsilon_1,\ \upsilon_2,\ \ldots,\ \upsilon_{t_u})]$$

is called the *Weyr characteristic of A.*

9 Theorem. *Let A be a complex square matrix of order n, let $\lambda_1, \lambda_2, \ldots, \lambda_u$ be its mutually distinct eigenvalues and let*

$$\eta(A) = [(\alpha_1,\ \alpha_2,\ \ldots,\ \alpha_{t_1}),\ (\beta_1,\ \beta_2,\ \ldots,\ \beta_{t_2}),\ \ \ldots,\ (\upsilon_1,\ \upsilon_2,\ \ldots,\ \upsilon_{t_u})]$$

be the Weyr characteristic of A. Then the sum of all characteristic numbers of the Weyr characteristic of A is equal to the order n of A.

The sum of the characteristic numbers associated with the eigenvalue λ_i is equal to its multiplicity s_i, and the sum of all multiplicities of the mutually distinct eigenvalues is equal to order n of A. Thus,

$$n = \sum_{i=1}^{u} s_i,$$

where

$$s_1 = \sum_{i=1}^{t_1} \alpha_i, \qquad s_2 = \sum_{i=1}^{t_2} \beta_i, \qquad \ldots, \qquad s_u = \sum_{i=1}^{t_u} \upsilon_i.$$

10 Theorem. *Two complex matrices are similar if and only if they have the same distinct eigenvalues and the associated Weyr characteristics.*

11 Theorem. *Let A be a square complex matrix, let $\lambda_1, \lambda_2, \ldots, \lambda_u$ be its mutually distinct eigenvalues and let $t_1,\ t_2,\ \ldots,\ t_u$ be the numbers of their characteristic numbers (i.e. the indices of A associated with $\lambda_1, \lambda_2, \ldots, \lambda_u$). Then*

$$(\lambda - \lambda_1)^{t_1} (\lambda - \lambda_2)^{t_2}\ \ldots\ (\lambda - \lambda_u)^{t_u}$$

is the minimal polynomial of A.

12 Definition. Let A be a square complex matrix, let J be its Jordan canonical form and let λ be its eigenvalue. The nonincreasing sequence $\xi(\lambda) = (\xi_1, \xi_2, \ldots, \xi_q)$ of orders of the Jordan blocks associated with the eigenvalue λ of A is called the *Segre characteristic of A associated with the eigenvalue λ.*

13 Definition. Let $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_t)$ be a nonincreasing sequence of positive integers. The *Ferrers diagram corresponding to the sequence* α is defined as a diagram which is formed by t columns of dots, such that the jth column (from the left) has α_j dots.

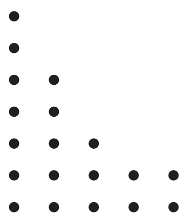
14 Definition. Let $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_t)$ be a nonincreasing sequence of positive integers. The sequence α^* is said to be *dual to* α if it is the sequence or row lengths of the Ferrers diagram corresponding to α (read upwards).

15 Remark. The dual sequence is often referred to as the conjugate sequence.

16 Theorem. *The Weyr characteristic of A associated with the eigenvalue λ is the dual sequence to the Segre characteristic of A associated with the same eigenvalue λ.*

Hence, the first characteristic number η_1 of the Weyr characteristic $(\eta_1, \eta_2, \ldots, \eta_t)$ of A associated with the eigenvalue λ is the number of blocks in the Jordan canonical form associated with λ. The number η_2 is the number of blocks in the Jordan canonical form that have order at least 2, the number η_3 is the number of blocks in the Jordan canonical form that have order at least 3, etc.

For example, if $\xi(\lambda) = (7, 5, 3, 2, 2)$, then the corresponding Ferrers diagram is



and $\eta(\lambda) = (5, 5, 3, 2, 2, 1, 1)$.

17 Definition. Let A be a square complex matrix, let λ be its eigenvalue. The *generalized eigenspace* $\mathrm{GKer}\,(A - \lambda E)$ of A associated with λ is the set $\mathrm{Ker}\,(A - \lambda E)^t$, where t is the index of A associated with λ.

Clearly $\mathrm{GKer}\,(A - \lambda E) = \mathrm{Ker}\,(A - \lambda E)^n$.

The multiplicity of λ is equal to the number of the dots in the corresponding Ferrers diagram and it is also equal to the dimension of $\mathrm{GKer}\,(A - \lambda E)$.

For $i \geq j$, the notation ${}_{i\times j}E$ denotes the matrix with i rows and j columns in which the first i rows form the identity matrix and the remaining $i - j$ rows consist of zeros.

$$
{}_{i\times j}E = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 1 \\ 0 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 0 \end{pmatrix}
\begin{matrix} \left.\vphantom{\begin{matrix}1\\0\\ \dots \\ 0\end{matrix}}\right\} j \\ \left.\vphantom{\begin{matrix}0\\ \dots \\ 0\end{matrix}}\right\} i-j \end{matrix}
$$

18 Definition. Let A be a square complex matrix, let λ be its eigenvalue and let $\eta(\lambda) = (\eta_1, \eta_2, \dots, \eta_t)$ be the Weyr characteristic of A associated with λ. The *Weyr block of A associated with λ* is defined as the matrix of the form

$$
\begin{pmatrix}
\begin{matrix} \lambda & 0 & \dots & 0 \\ 0 & \lambda & \dots & 0 \\ \vdots & \vdots & \ddots & 0 \\ 0 & 0 & \dots & \lambda \end{matrix} & {}_{\eta_1\times\eta_2}E & O & \cdots & O \\
O & \begin{matrix} \lambda & 0 & \dots & 0 \\ 0 & \lambda & \dots & 0 \\ \vdots & \vdots & \ddots & 0 \\ 0 & 0 & \dots & \lambda \end{matrix} & {}_{\eta_2\times\eta_3}E & \cdots & O \\
O & O & \begin{matrix} \lambda & 0 & \dots & 0 \\ 0 & \lambda & \dots & 0 \\ \vdots & \vdots & \ddots & 0 \\ 0 & 0 & \dots & \lambda \end{matrix} & \ddots & O \\
\vdots & \vdots & \vdots & \ddots & {}_{\eta_{t-1}\times\eta_t}E \\
O & O & O & \cdots & \begin{matrix} \lambda & 0 & \dots & 0 \\ 0 & \lambda & \dots & 0 \\ \vdots & \vdots & \ddots & 0 \\ 0 & 0 & \dots & \lambda \end{matrix}
\end{pmatrix}
$$

where O is the zero matrix.

19 Definition. Let A be a complex square matrix. The block matrix whose diagonal blocks are the Weyr blocks associated with all distinct eigenvalues of A is called the *Weyr canonical form of A*.

For example, if the eigenvalues of a matrix A are the numbers 2 and -3 and if the Weyr characteristic of A associated with the eigenvalue 2 is

$\eta(2) = (3,\, 2,\, 1,\, 1)$ and the Weyr characteristic of A associated with the eigenvalue -3 is $\eta(-3) = (3,\, 2)$, then the Weyr canonical form W of A is

$$W = \left(\begin{array}{ccc|cc|c|c|ccc|cc}
2 & & & 1 & & & & & & & & \\
 & 2 & & & 1 & & & & & & & \\
 & & 2 & & & & & & & & & \\ \hline
 & & & 2 & & 1 & & & & & & \\
 & & & & 2 & & & & & & & \\ \hline
 & & & & & 2 & 1 & & & & & \\ \hline
 & & & & & & 2 & & & & & \\ \hline
 & & & & & & & -3 & & & 1 & \\
 & & & & & & & & -3 & & & 1 \\
 & & & & & & & & & -3 & & \\ \hline
 & & & & & & & & & & -3 & \\
 & & & & & & & & & & & -3
\end{array}\right).$$

20 Theorem. *Up to permutation of the Weyr blocks, each square complex matrix A is similar to a unique Weyr canonical form.*

* * *

Responses to Weyr theory abroad

Despite the fact that several of Eduard Weyr's results advanced the foreign knowledge, he practically gained no recognition abroad for about a century.[3]

Weyr is cited in several articles, monographs or encyclopaedias. However, there were only sporadic acknowledgements in few publications until the 1980s. On top of it the Weyr characteristic was often recognised only as a dual sequence to the Segre characteristic. Weyr's name in association with canonical form of matrices was almost never mentioned. From the 8th decade of the 20th century this situation began to change. Group of mathematicians close to Hans Schneider and Daniel Hershkowitz published a significant number of articles devoted to the Weyr characteristic, defined in terms of matrix theory, and various other characteristics introduced in terms of graph theory. Works concerning these problems have been published even in the 3rd millennium.

Since the 1980s, the Weyr characteristic in connection with the so-called matrix pencils began to be suitably used by mathematicians belonging to Greek school (Nicos Karcanias, Grigoris Kalogeropoulos, Panayiotis Psarrakos, etc.). Then followed the reactions of Spanish mathematicians (Juan Miguel Gracia, Inmaculada de Hoyos, Ion Zaballa, María Asunción Beitia, Francisco Enrique Velasco, etc.), who studied problems closely linked with the so-called control

[3] During Weyr's life, his works were appreciated by James Joseph Sylvester, who was then an authority in the worldwide mathematical community.

theory. On the turn of the 20th and 21st century the Weyr characteristic was also recognised by mathematicians working in Sweden (Erik Elmroth, Bo Kågström, Pedher Johansson, Stefan Johansson, etc.), who used it in the stratification of orbits and bundles of matrices and matrix pencils).

Further awareness of the Weyr characteristic can be merited to Helene Shapiro who in 1999 published the article *The Weyr characteristic* [Sh2]. This gave the basic knowledge of the Weyr characteristic and Weyr canonical form in a simple and intelligible manner. The work was published in *The American Mathematical Monthly* magazine, which is intended for a wider mathematical community (not just experts in linear algebra), therefore the fundamental facts of Weyr theory became known to a relatively large number of readers.

The Weyr canonical form can be found in literature since the 1980s. However, it does not reflect the acknowledgements of Weyr's results in the real sense. The mathematicians working with this form rediscovered it themselves while searching for a new form of matrices with certain properties. They often never heard of Eduard Weyr before. Personalities most responsible for the revival of the Weyr canonical form are Genrich Ruvimovich Belitskii, Vladimir Vasil'evich Sergeichuk, Junzo Watanabe, Tadahito Harima, Kevin C. O'Meara, Charles Irvin Vinsonhaler and Helene Shapiro. It was probably the last named mathematician who used the historically correct term, i.e. *Weyr canonical form* for the first time in a published text. It was in the above mentioned article *The Weyr characteristic* in 1999. Vladimir V. Sergeichuk became aware of the connection of the Weyr characteristic and Weyr canonical form even before. Prior to 1999 this term was used under different names, such as the *second Jordan canonical form*, *modified Jordan matrix*, *H-form*, etc. Nowadays, more than one century since Weyr's death, the Weyr characteristic and the Weyr canonical form occur rather frequently in contemporary foreign language works. In 2013, Roger Alan Horn and Charles R. Johnson published second edition of their famous monograph *Matrix Analysis* [HJ1], which appeared in 1985 for the first time. Roger A. Horn rewrote one chapter dealing with the canonical forms, and recast the exposition in terms of the Weyr characteristic. The Weyr canonical form has become much better known in the last few years and even the monograph *Advanced Topics in Linear Algebra: Weaving Matrix Problems through the Weyr Form* [OCV1] devoted to this topic was published by Kevin C. O'Meara, John Clark a Charles I. Vinsonhaler in 2011.

Although Weyr's outcomes were almost forgotten for approximately a century, their recognitions in the last decades affect various and very specialized branches of mathematics. It is therefore impossible to present expert contents in brief. Let us mention at least the elements of one of the disciplines which operates with Weyr characteristic and whose results were published in great deal. It concerns the above mentioned works of Hans Schneider, Daniel Hershkowitz and their co-authors who studied a close connection between the Weyr characteristic and certain sequences formulated in terms of graph theory.

1 Remark. We will denote the spectrum of a square matrix A, i.e. the set of all eigenvalues of the matrix A, by $\sigma(A)$.

2 Definition. The *spectral radius* $\varrho(A)$ of a square matrix A is defined to be

$$\varrho(A) = \max\{|\lambda|\,;\lambda \in \sigma(A)\}\,.$$

3 Definition. Let $A = (a_{ij})$ be a square complex matrix or order n. Then A is called *reducible* if there exists a permutation matrix P such that PAP^T has the form

$$\begin{pmatrix} K & O \\ L & M \end{pmatrix},$$

where K and L are square matrices of order at least 1 and O is the zero matrix. If no such permutation exists, then A is said to be *irreducible.* In other words, a square complex matrix is reducible if it can be transformed into the above-mentioned form by the same permutation of rows and columns. If A is a square complex matrix of order 1, then A is irreducible.

4 Definition. A matrix A is said to be *nonnegative* if all of its entries are nonnegative. A matrix A is said to be *positive* if all of its entries are positive.

5 Definition. A square matrix A is called an *M-matrix* if there is a nonnegative square matrix B and a number $k \geq \varrho(B)$ such that $A = kE - B$.

The development of the Perron-Frobenius spectral theory (1907, 1912) of nonnegative matrices was followed by an intensive study of the interrelation between the spectral properties of matrices (for example, the Weyr characteristic) and certain graph theoretic properties (for example, the level characteristic). The most important results of this type are formulated in the following theorems.

6 Theorem. *Let A be a square nonnegative matrix. Then the spectral radius $\varrho(A)$ of A is itself an eigenvalue of A. Futhermore, if A is an irreducible matrix, then $\varrho(A)$ is a simple eigenvalue of A and there is a positive eigenvector associated with the eigenvalue $\varrho(A)$.*

Some results of the Perron-Frobenius theory can be extended to singular M-matrices.

7 Theorem. *A singular M-matrix A has a nonnegative nullvector v, i.e. a non-zero vector v which satisfies $Av^T = o^T$. Furthermore, if A is irreducible, then* 0 *is a simple eigenvalue of A with an associated positive eigenvector.*

It is natural to study properties of the eigenvalue $\varrho(B)$ of a nonnegative matrix B. Clearly, an M-matrix $A = kE - B$ is singular if and only if $\varrho(B) = k$. It is also well known that a matrix B has an eigenvalue λ of multiplicity s if and only if the matrix $\lambda E - B$ has the eigenvalue 0 of multiplicity s. For convenience, results are usually formulated in an equivalent form in terms of a singular M-matrix with respect to the eigenvalue 0. Thus, we may equivalently study the zero eigenvalue of a singular M-matrix $A = \varrho(B)E - B$ instead of studying the spectral radius $\varrho(B)$ of a nonnegative matrix B.

8 Definition. Let A be a square matrix of order n. The *graph* $G(A)$ of A is the graph with vertices 1, 2, ..., n which has an arc from i to j if and only if $a_{ij} \neq 0$.

It is well known that a square nonnegative matrix is irreducible if and only if its associated directed graph is strongly connected.

9 Remark. All the graphs in this summary (and also in the monograph) are simple directed graphs.

10 Definition. Let A be a block matrix of the form

$$\begin{pmatrix} A_{11} & A_{12} & \cdots & A_{1p} \\ A_{21} & A_{22} & \cdots & A_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ A_{p1} & A_{p2} & \dots & A_{pp} \end{pmatrix}.$$

The *reduced graph* $R(A)$ of A is defined to be the directed graph with vertices 1, 2, ..., p and such that there is an arc from i to j if and only if $A_{ij} \neq O$.

11 Definition. Let A be a matrix in a $p \times p$ form whose every diagonal block A_{ii}, $i = 1, 2, \ldots, p$, is square. A vertex i in the reduced graph $R(A)$ is called *singular* if the matrix A_{ii} is singular.

12 Definition. The *path* in a graph is a sequence $(v_1, v_2, \ldots, v_m)$ of distinct vertices such that each two consecutive vertices v_k and v_{k+1}, $k = 1, 2, \ldots, v_{m-1}$, are joined by an arc in the graph. We also consider one vertex to be a path. The *length of a path* is the number of vertices on this path.

13 Definition. Let A be a block matrix of the form

$$\begin{pmatrix} A_{11} & O & \cdots & O \\ A_{21} & A_{22} & \cdots & O \\ \vdots & \vdots & \ddots & \vdots \\ A_{q1} & A_{q2} & \dots & A_{qq} \end{pmatrix}.$$

If all blocks which lie on the main diagonal are irreducible square matrices, then we say that this form is the *Frobenius normal form*.

It is well known that every square matrix can be transformed by the same permutation of rows and columns (i.e. by permutation similarity) to the Frobenius normal form.

14 Definition. Let $R(A)$ be the reduced graph of a square matrix A and let i be a singular vertex in $R(A)$. The *level* of the singular vertex i in $R(A)$ is the maximal number of singular vertices on a path in $R(A)$ that terminates at the vertex i.

15 Definition. Let m be the maximal level of a singular vertex in the reduced graph $R(A)$ of a square matrix A. We say that the sequence $\lambda(A) = (\lambda_1, \lambda_2, \ldots, \lambda_m)$ is the *level characteristic of* A if λ_k is the number of singular vertices of $R(A)$ of level k.

We shall always assume that A is a square matrix of order n and that A is in the Frobenius normal form. Then the level characteristic is uniquely determined.

Suppose we wish to determine the level characteristic of the real matrix

$$A = \left(\begin{array}{c|cc|c|c|c} 0 & 0 & 0 & 0 & 0 & 0 \\ \hline 0 & 2 & -3 & 0 & 0 & 0 \\ 1 & -4 & 6 & 0 & 0 & 0 \\ \hline 2 & 0 & 2 & 2 & 0 & 0 \\ \hline 1 & 0 & 0 & -3 & 0 & 0 \\ \hline 3 & 0 & 0 & -2 & 0 & 0 \end{array}\right).$$

The reduced graph of A is given in Figure 1.

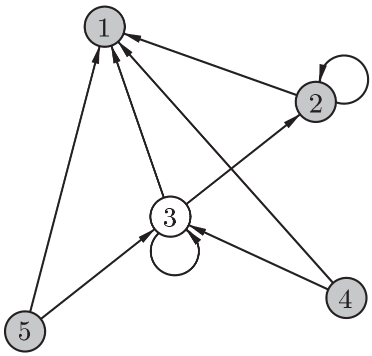


Figure 1

In figures, the singular vertices have been highlighted. The paths (5, 3, 2, 1) and (4, 3, 2, 1) terminate at vertex 1 and contain the maximal number of singular vertices (vertices 5, 2, 1 and 4, 2, 1) among all paths in $R(A)$. The level of the singular vertex 1 is three. There are two paths in $R(A)$ which terminate at vertex 2. Both of them include two singular vertices. Thus, the level of the singular vertex 2 is two. The level of the singular vertex 4 is one. The same is true of the singular vertex 5. We have $\lambda(A) = (2, 1, 1)$.

16 Definition. For a singular matrix A of order n let t be the maximal positive integer such that $\text{null}\, A^{t-1} < \text{null}\, A^t$. The sequence $\eta(A) = (\eta_1, \eta_2, \ldots, \eta_t)$ is called the *height characteristic* of A if $\eta_k = \text{null}\, A^k - \text{null}\, A^{k-1}$.

The height characteristic of A agrees with the Weyr characteristic of A associated with the eigenvalue 0.

One may ask about the connection between the height characteristic (or equivalently the Segre characteristic) associated with the eigenvalue 0 of a singular M-matrix and its associated reduced graph.

Research in this direction has a long tradition and originated in the early work of Hans Schneider at the beginning of the second half of the 20th century. See, for example, Schneider's article *The elementary divisors associated with* 0*, of a singular M-matrix* [Sc2], which was published in 1956. Authors of the early works did not explicitly use graphs. They mentioned neither the height characteristic nor the level characteristic.

Progress was made in the 1970s and 1980s when people began to use graph-theoretic methods. See, for example, the article *On the singular graph and the Weyr characteristic of an M-matrix* [RS1], which was written by Daniel James Richman and Hans Schneider in 1978.

There are two extreme cases in which the height characteristic is equal to the level characteristic. The next two theorems were proved by Schneider in his thesis *Matrices with non-negative elements* in 1952 without using a graph.

17 Theorem. *Let A be an M-matrix. Then the following statements are equivalent:*

(i) *The height characteristic* $\eta(A)$ *of* A *is* (t) (*or, equivalently, the Segre characteristic is* $(1, 1, \ldots, 1)$).

(ii) *The level characteristic* $\lambda(A)$ *of* A *is* (t).

Thus, in the case when each path of $R(A)$ has at most one singular vertex we have

$$\lambda(A) = \eta(A).$$

18 Example. Consider

$$A = \begin{pmatrix} a & 0 & 0 & 0 & 0 \\ -b & 0 & 0 & 0 & 0 \\ -c & 0 & d & 0 & 0 \\ 0 & 0 & -e & 0 & 0 \\ 0 & 0 & -f & 0 & 0 \end{pmatrix},$$

where a, b, c, d, e, f are positive numbers. The matrix A is an M-matrix, it is in the Frobenius normal form, the blocks on the main diagonal are matrices of order 1. The reduced graph of A agrees with the graph of A and is shown in Figure 2.

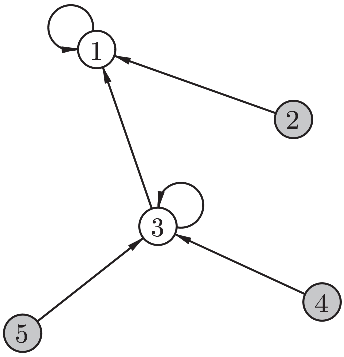


Figure 2

The levels of all singular vertices of the reduced graph $R(A)$ are one and we have $\lambda(A) = (3)$. Let's compute the powers of matrices A, A^2, ... and determine their nullities: $\operatorname{null} A = \operatorname{null} A^2 = 3$. We have $\eta(A) = (3)$ and $\lambda(A) = \eta(A)$.

19 Theorem. *Let A be an M-matrix. Then the following statements are equivalent:*

(i) *The height characteristic* $\eta(A)$ *of* A *is* $(1, 1, \ldots, 1)$ (*or, equivalently, the Segre characteristic is* (t)).
(ii) *The level characteristic* $\lambda(A)$ *of* A *is* $(1, 1, \ldots, 1)$.

We have

$$\lambda(A) = \eta(A)$$

again. This is the case when all singular vertices of the reduced graph $R(A)$ of A lie on a path.

20 Example. Let

$$A = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & a & 0 & 0 \\ -b & -c & 0 & 0 \\ -d & 0 & -e & 0 \end{pmatrix},$$

where a, b, c, d, e, f are positive numbers. The matrix A is an M-matrix, it is in the Frobenius normal form, the blocks on the main diagonal are matrices of order 1. The reduced graph of A agrees with the graph of A and is given in Figure 3.

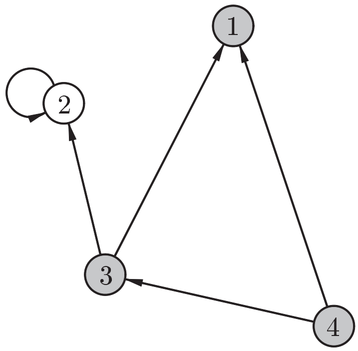


Figure 3

All singular vertices lie on the same path. The level of the vertex 4 is one, the level of the vertex 3 is two and the level of the vertex 1 is three. Thus, $\lambda(A) = (1, 1, 1)$. Because $\operatorname{null} A = 1$, $\operatorname{null} A^2 = 2$, $\operatorname{null} A^3 = \operatorname{null} A^4 = 3$, we have $\eta(A) = (1, 1, 1)$ and $\lambda(A) = \eta(A)$.

This equality does not hold in general. For example, the height characteristic of the following matrix is not equal to its level characteristic.

21 Example. Consider

$$A = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ -a & -b & c & 0 & 0 \\ -d & -e & -f & 0 & 0 \\ 0 & 0 & 0 & -g & 0 \end{pmatrix},$$

where a, b, c, d, e, f, g are positive integer. The reduced graph of A agrees with the graph of A and is given in Figure 4.

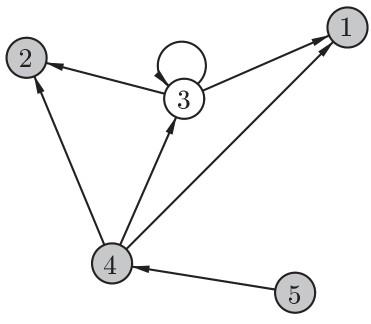


Figure 4

The level characteristic $\lambda(A)$ is $(1, 1, 2)$ and the height characteristic $\eta(A)$ is $(2, 1, 1)$.

It is natural to ask about the relationship between the level characteristic $\lambda(A)$ and the height characteristic $\eta(A)$, where A is an M-matrix.

In 1989, Hans Schneider and Daniel Hershkowitz introduced the concepts of height basis, level basis, and height-level basis for a generalized nullspace of an M-matrix A in their article *Height bases, level bases, and the equality of the height and the level characteristics of an M-matrix* [HS3]. Then they gave twelve conditions equivalent to $\lambda(A) = \eta(A)$, most of them in terms of above-mentioned basis. They added another twenty three equivalent conditions in their article *Combinatorial bases, derived Jordan sets and the equality of the height and level characteristic of an M-matrix* [HS4], which was published in 1991.

22 Definition. Let $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_t)$ and $\beta = (\beta_1, \beta_2, \ldots, \beta_t)$ be sequences of nonnegative integers with the same number t of elements. We say that β *majorizes* α if

$$\alpha_1 + \alpha_2 + \ldots + \alpha_t \leq \beta_1 + \beta_2 + \ldots + \beta_t \qquad \text{for all} \quad k < t$$

and

$$\alpha_1 + \alpha_2 + \ldots + \alpha_t = \beta_1 + \beta_2 + \ldots + \beta_t.$$

We denote this relation by $\alpha \preceq \beta$.

23 Theorem. *For an M-matrix A we have*

$$\lambda(A) \preceq \eta(A).$$

We focused on M-matrices (or, equivalently, on nonnegative matrices) because their spectral and graph properties were studied for about seventy years after the Perron-Frobenius theory was published. Other classes of matrices were investigated later. The expansion of the study to more general matrices resulted in many theorems during the last 25 years.

We formulate at least the following one which was proved in the article *A majorization relation between the height and the level characteristics* [He1], published by Daniel Hershkowitz in 1989.

24 Theorem. *Let A be a block triangular matrix with square diagonal blocks such that* 0 *is a simple eigenvalue of the singular diagonal blocks. Let* $\hat{\lambda}$ *be the level characteristic of A reordered in a nonincreasing order. Then*

$$\hat{\lambda}(A) \preceq \eta(A).$$

Because the length of this summary is limited, we restricted our discussion to the level characteristic only. However, for certain classes of matrices, the level characteristic can be replaced by other sequences denoted by $\delta(G(A))$ and $\pi(G(A))^*$. These sequences are majorized by the height characteristic and they simultaneously majorize the level characteristic.

* * *

Conclusion

Instead of writing a "typical conclusion", let us quote Rogen Alan Horn, the author of the monograph *Matrix Analysis*:

Weyr's 1890 paper in vol. 1 of Monatshefte Math. Physik ... is astonishingly modern in its notation and spirit. It could be read and understood by any (German-speaking) student today who has had a first course in linear algebra. In contrast, the 1870 book of C. Jordan ... that contains his eponymous canonical form is unintelligible to a modern reader; no one thinks of linear algebra in terms of "substitutions" any more. I hope that the second edition of Matrix Analysis will help many students and researchers to learn a little about this great man's contributions to mathematics.