

The Order of a Singularity in Fuchs' Theory

By

JÜRGEN MOSEK

I. Introduction

1. This paper is concerned with systems of ordinary differential equations

$$(1) \quad \frac{dy}{dx} = A(x)y$$

where y is a vector with n components and $A(x)$ an n by n matrix. The elements of $A(x)$ are assumed to be analytic functions of the complex variable x for $0 < |x| < \varrho$ with a pole at $x=0$:

$$(2) \quad A(x) = x^{-p} \sum_{\nu=0}^{\infty} A_{\nu} x^{\nu} \quad (A_0 \neq 0)$$

where p is an integer and the A_{ν} constant matrices. The aim of Fuchs' theory is to study the nature of the solution vectors $y(x)$ near the singularity $x=0$.

In most treatments of this problem one restricts the attention to the case where A_0 is a matrix with distinct eigenvalues in which case one can describe the solution by asymptotic series for $p > 1$. For $p=1$ one has a regular singularity, i.e. all solutions $y(x)$ of (1) grow at most like a finite power of $|x|$

$$(3) \quad |y| \leq c|x|^N$$

for small x and with some norm $|y|^1$.

For the degenerate case, however, where A_0 has several zero eigenvalues or even is nilpotent this problem is very complicated and has not been treated in this generality. It can happen, for instance, that even in the case $p > 1$ one has a regular singularity (see example below). On the other hand it was proven by HORN [1] that in case $x=0$ is a regular singularity one can transform the system (1) by

$$(4) \quad y = T(x)z$$

into a system

$$(5) \quad \frac{dz}{dx} = B(x)z, \quad B = T^{-1}AT - T^{-1} \frac{d}{dx} T$$

where $B(x)$ has a pole of first order only. Here $T(x)$ is a matrix whose elements are analytic in some neighborhood $0 < |x| < \varrho$ with at most a pole at $x=0$, and which satisfies $\det T(x) \neq 0$.

¹⁾ Since $y(x)$ is multivalued in general one has to restrict the argument of x to a fixed sector $|\arg x| < \text{const.}$

We will call systems (4) and (5) which are related by such a transformation (4) "equivalent", since the behavior of the solutions near a singularity is essentially the same, namely up to functions which have a pole near the origin. Here we require only that $\det T(x)$ is not identically zero but allow that $\det T(0)$ vanishes²⁾.

Thus Horn's theorem can be expressed as follows: A system (1) has a regular singularity at $x=0$ if and only if it is equivalent to a system (2) with a pole of first order. Even though this criterion gives a necessary and sufficient condition for a regular singularity it cannot be used to decide about the nature of the singularity for a preassigned system. There is no method of constructing the transformation $T(x)$, and it is not even clear, how many terms of the transformation matrix

$$T(x) = \sum_{\nu=-N}^M x^{\nu} T_{\nu}$$

are necessary for the reduction. It is the purpose of this note to give a criterion which allows the decision of this question in a definite number of steps.

2. For any system of the form (1) we define the rational number

$$(6) \quad m(A) = p - 1 + \frac{r}{n} \geq 0$$

as the *order* of A where $r=r(A_0)$ is the algebraic rank of A_0 , $0 < r \leq n$. If $p - 1 + \frac{r}{n} < 0$ we set $m(A) = 0$. For every system (1) we introduce the number

$$(6') \quad \mu(A) = \min_T m(T^{-1}AT - T^{-1}T')$$

i.e. $\mu(A)$ is the minimum value of the order of all systems which are equivalent to (1). One can consider $\mu(A) - 1$ as a generalization of "Poincaré's rank" of a singularity. The matrix $A(x)$ is called *reducible*, if $m(A) > \mu(A)$.

Before stating the main results we illustrate the significance of $\mu(A)$ for the nilpotent matrix

$$A(x) = \begin{pmatrix} x^{-q} & -x^{-2q} \\ 1 & -x^{-q} \end{pmatrix}$$

where q is a positive integer. Obviously $m(A) = 2q - \frac{1}{2}$ but this matrix is reducible. With the transformation matrix

$$T(x) = \begin{pmatrix} 1 & 0 \\ x^q & x^{q+\gamma} \end{pmatrix}, \quad \gamma = \left[\frac{q}{2} \right]^3$$

one computes $B = T^{-1}AT - T^{-1}T'$ to

$$B(x) = \begin{pmatrix} 0 & -x^{-q+\gamma} \\ -qx^{-\gamma-1} & -(q+\gamma)x^{-1} \end{pmatrix}.$$

²⁾ A similar concept of equivalence was introduced by G. D. BIRKHOFF [2], who required that $T(0)$ is the identity matrix. In this case, A_0 is unchanged, and the order p of the pole is the same for equivalent systems.

³⁾ $[x]$ stands for the greatest integer $\leq x$.

One can show that $B(x)$ is not reducible and one finds

$$\mu(A) = m(B) = \frac{q+1}{2}.$$

In particular the minimum order of the pole $x=0$ is

$$\gamma + 1 = \left[\frac{q}{2} \right] + 1.$$

In this simple example the order $m(A)$ can be reduced from $2q - \frac{1}{2}$ to $\frac{1}{2}(q+1)$.

The regular singularity is characterized by the inequality

$$\mu \leq 1$$

according to Horn's theorem. In the following we will assume that

$$m(A) > 1$$

and investigate under which conditions $A(x)$ is reducible.

The reason for the above refined definition of the order $m(A)$ of the singularity lies in the fact that it is possible to give a necessary and sufficient criterion for reducibility which involves the first two coefficients A_0, A_1 of $A(x)$ only. This statement is contained in

THEOREM 1. *If $m(A) > 1$ the system (1) is reducible if and only if the polynomial*

$$(7) \quad \mathfrak{P}(\lambda) = x^r \det(\lambda I + x^{p-1} A(x))|_{x=0}, \quad r = r(A_0)$$

vanishes identically in λ .

The polynomial $\mathfrak{P}(\lambda)$ depends on A_0, A_1 only, since in

$$x^{p-1} A(x) = \frac{A_0}{x} + A_1 + x A_2 + \dots$$

A_0 has the rank r . Hence in forming the determinant in (7) one has to take r columns from A_0 and the others from $\lambda I + A_1 + x A_2 + \dots|_{x=0} = \lambda I + A_1$. Thus

$$\mathfrak{P}(\lambda) = \sum_{\nu=0}^{n-r} \lambda^\nu \mathfrak{P}_\nu(A_0, A_1)$$

where the coefficients $\mathfrak{P}_\nu(A_0, A_1)$ are homogeneous polynomials in the coefficients of A_0, A_1 of degree $n - \nu$. Therefore the above criterion requires

$$\mathfrak{P}_\nu(A_0, A_1) = 0 \quad \text{for } \nu = 0, \dots, n - r.$$

Since reducibility depends on the coefficients A_0, A_1 one can expect that the reduction of $m(A)$ can be achieved by simple transformation matrices $T(x)$. This is expressed by

THEOREM 2. If $A(x)$ is reducible and $m(A) > 1$ then the reduction can be carried out with a matrix T of the form

$$T = (P_0 + xP_1) \text{diag}(1, 1, \dots, 1, x, x, \dots, x)$$

where P_0, P_1 are constant matrices and $\det P_0 \neq 0$.

REMARK. This theorem expresses only that $B = T^{-1}AT - T^{-1}T'$ satisfies $m(B) < m(A)$, not that $m(B)$ equals $\mu(A)$.

Theorem 2 limits the computation of T to two matrices. After this reduction has been carried out one can apply Theorem 1 to check whether a further reduction is possible. Here it is important to observe that $\mathfrak{P}(\lambda)$ will change from one step to the next since B is not obtained by a similarity transformation but by $B = T^{-1}AT - T^{-1}T'$. If the system possesses a regular singularity one has to reduce the order from $m(A)$ to $\mu = 1$ which requires at most $(p-2)n + r = n(m(A) - 1)$ steps. In this sense one can consider Theorem 1 as a criterion for a regular singularity, although it is not given in "closed form". The same remark refers more generally to the computation of $\mu(A)$. After one has reached the last step for which $m(A) = \mu(A)$ one has $\mathfrak{P}(\lambda) \neq 0$ which is the characteristic property of A_0, A_1 in the irreducible case.

3. It is interesting to note that for ordinary differential equation

$$(8) \quad \frac{d^n u}{dx^n} + a_1(x) \frac{d^{n-1} u}{dx^{n-1}} + \dots + a_{n-1}(x) \frac{du}{dx} + a_n(x) u = 0$$

the invariant μ can be computed explicitly and not only recursively. Here the $a_k(x)$ are assumed to be analytic in $0 < |x| < \rho$ with at most a pole at $x = 0$. The equation (8) can be rewritten as a first order system (1) for which the number $\mu = \mu(A)$ has been defined.

This number can be computed as follows: Let $\lambda_k \geq 0$ be the smallest integer such that

$$x^{\lambda_k} a_k(x) \text{ is regular at } x = 0.$$

THEOREM 3. With the above definition let p be the smallest integer with

$$p \geq \frac{\lambda_k}{k}, \quad r = \text{Max}_k (\lambda_k - (p-1)k) > 0.$$

Then

$$(9) \quad \mu = p - 1 + \frac{r}{n}$$

is the required invariant provided $p > 1$. If $p \leq 1$ one has $\mu \leq 1$.

This theorem can be considered as a generalization of Fuchs' theorem which states that (8) has a regular singularity if and only if $\lambda_k \leq k$, i.e. $p \leq 1$. Such a criterion is not known for systems (1). A similar situation occurs in the computation of μ , which can be done explicitly for differential equations (8) while for systems only a "recursive" criterion can be given.

The differential equations (8) can be chosen as models to exhibit all non-negative rational numbers with denominator n as possible values for μ . For $0 < \mu \leq 1$ one has a regular singularity and for $\mu = 0$ a regular point.

If the matrix $A(x)$ is real and symmetric or hermitean the computation of μ can be given explicitly since only the reduced case occurs.

4. The above question is closely related to the problem of invariants of $A(x)$ under the equivalence

$$(10) \quad A \sim B = T^{-1} A T - T^{-1} T'$$

where $T(x)$ is a matrix described under (4). One might expect similar normal forms as they are known under similarity transformations. That the situation is quite different in both cases is seen from the fact that the invariants under similarity consist of at least n power series (for instance the coefficients of the characteristic polynomial) while the equivalence (10) admits only finitely many numbers as invariants for A . This follows from a theorem of G. D. BIRKHOFF [2] who proved that one can find a $T(x)$ with $T(0) = I$ such that

$$B = \sum_{\nu=1}^p B_{p-\nu} x^{-\nu}$$

does not contain x in positive powers. His method requires solving an integral equation and might be called *transcendental*, since it requires the knowledge of infinitely many terms in A . In contradistinction to Birkhoff's result this paper is concerned with terms of negative exponents in x and an *algebraic* description of those terms. While in Birkhoff's paper the term $T^{-1} T'$ in (10) is to be considered as the principal part $T^{-1} A T$ is the main term in this note.

The equivalence relation (10) has been investigated in several papers, in particular, by LOEWY [6], where the question of decomposability of a matrix or a matrix complex is treated. However, these results are not used and not needed for the proof of the above results.

In this paper we consider the coefficients of all analytic functions to be complex numbers although it is clear that the proofs can be generalized to commutative fields without zero divisors.

II. 3 Lemmata

The proof of the above theorem is based on 3 lemmata which are proven in this section. The first lemma can be considered as a generalization of the fact that an analytic function $f(z) \equiv 0$ which has at most a pole at $z=0$ can be written in the form $f(z) = z^\alpha g(z)$ where $g(z)$ is analytic at $z=0$ and $g(0) \neq 0$. The number α will be replaced by a diagonal matrices in the corresponding statement for matrices.

LEMMA 1. Let

$$T(x) = \sum_v T_v x^v$$

be a sum with finitely many negative v and $\det T(x) \not\equiv 0$. Then $T(x)$ can be represented in the form

$$T(x) = P(x) x^\alpha Q(x)$$

where $P(x)$ is a polynomial with $\det P(x) \equiv 1$ and $Q(x)$ a power series with $\det Q(0) \neq 0$ and

$$x^\alpha = \text{diag}(x^{\alpha_1}, \dots, x^{\alpha_n})$$

where $\alpha_1 \leq \alpha_2 \leq \dots \leq \alpha_n$ are integers.

REMARK 1. $P(x)$ is a unit in the ring of polynomials since on account of $\det P(x) \equiv 1$ the inverse $P^{-1}(x)$ also is a polynomial. $Q(x)$ is a unit in the ring of power series while x^α represents the essential part of $T(x)$.

REMARK 2. Like in Lemma 1 one can find a representation of the form

$$T(x) = Q(x) x^\alpha P(x)$$

with the same properties specified for P , Q in Lemma 1.

PROOF. We prefer to prove the statement in the form of remark 2

$$T(x) = Q(x) x^\alpha P(x)$$

considering the column vectors $t_j(x)$ of $T(x)$ instead of the rows. The matrix $P(x)$ will be built up by multiplication out of "elementary" matrices $E(x)$, F :

$$(*) \quad E = I + E_{kl}(x) \quad (k \neq l)$$

where E_{kl} has all elements equal to zero but one element in the k -th row and l -th column which is any polynomial $p_{kl}(x)$, $k \neq l$. Obviously $\det E \equiv 1$. F stands for any permutation matrix

$$F = (\pm e_{j_1}, \pm e_{j_2}, \dots, \pm e_{j_n})$$

where e_l is the l -th unit vector and j_1, \dots, j_n is a permutation of $1, 2, \dots, n$. The signs \pm are arbitrary except for the condition $\det F = 1$. Thus F is independent of x .

Multiplying a matrix

$$T = (t_1, t_2, \dots, t_n)$$

with E from the right results in replacing t_l by $t_l + p_{kl}(x)t_k$ while multiplication with F from the right permutes the column vectors and changes the signs. Let P denote any product of several such matrices E and F . Obviously these matrices form a group all elements of which are matrices of polynomials with $\det P \equiv 1$.

To prove Lemma 1 we define the integers $\alpha_1, \alpha_2, \dots, \alpha_n$ successively by the following maximal conditions: Let α_1 be the greatest integer such that $x^{-\alpha_1} T(x)$ is a power series. The constant term of $x^{-\alpha_1} T(x)$ then is a nonzero matrix and applying a matrix F we can assume that

$$T(x)F = (x^{\alpha_1} s_1(x), x^{\alpha_2} s_2(x), \dots, x^{\alpha_n} s_n(x))$$

where $s_l(x)$ are power series and $s_1(0) \neq 0$. Obviously α_1 remains unchanged if T is replaced by TP .

α_2 is chosen as the greatest integer under all choices of P such that the representation

$$T(x)P = (x^{\alpha_1}r_1(x), x^{\alpha_2}r_2(x), x^{\alpha_3}r_3(x), \dots, x^{\alpha_n}r_n(x))$$

holds where the $r_l(x)$ are power series. Obviously $\alpha_2 \geq \alpha_1$. After having defined $\alpha_1, \dots, \alpha_{l-1}$ we choose α_l to be the greatest integer for which a representation

$$(2.1) \quad T(x)P(x) = (x^{\alpha_1}q_1(x), \dots, x^{\alpha_l}q_l(x), x^{\alpha_{l+1}}q_{l+1}(x), \dots, x^{\alpha_n}q_n(x))$$

holds. Again $\alpha_1 \leq \alpha_2 \leq \dots \leq \alpha_l$ is an obvious consequence of the definition. The existence of α_l follows from the assumption that $\det(T(x)P(x)) = \det T(x) \neq 0$, hence there is an integer γ with $T(x) = cx^\gamma + \dots, c \neq 0$. The estimate $\alpha_1 + \alpha_2 + \dots + \alpha_{l-1} + (n-l+1)\alpha_l \leq \gamma$ guarantees the existence of a maximal α_l .

Now it follows by induction that in any maximal representation (2.1)

$$(2.2) \quad \text{rank}(q_1(0), q_2(0), \dots, q_n(0)) \geq l.$$

For $l=1$ this is obvious. Assume this has been proven for $l-1$. If necessary one can achieve then that $q_1(0), q_2(0), \dots, q_{l-1}(0)$ are linearly independent vectors by applying a permutation matrix F to (2.1). If the statement (2.2) were wrong then

$$q_k(0) = \sum_{\lambda=1}^{l-1} c_{k\lambda} q_\lambda(0) \quad \text{for } k = l, \dots, n.$$

Applying a product of matrices E we can achieve

$$q_k(0) = 0 \quad \text{for } k = l, \dots, n$$

which implies that α_l can be increased. This contradiction proves (2.2). For $l=n$ we obtain

$$TP = (x^{\alpha_1}q_1(x), \dots, x^{\alpha_n}q_n(x)) = Q(x)x^\alpha$$

where

$$Q(x) = (q_1(x), \dots, q_n(x)), \quad x^\alpha = \text{diag}(x^{\alpha_1}, \dots, x^{\alpha_n})$$

is a power series with

$$\det Q(0) \neq 0.$$

Thus

$$T = Q(x)x^\alpha P^{-1}$$

and since P^{-1} also is a polynomial with $\det P \equiv 1$ Lemma 1 is proven in the form of remark 2. Applying the argument to the transposed of T one obtains the statement of the Lemma 1.

The representation of the Lemma 1 is not unique, the numbers $\alpha_1, \dots, \alpha_n$, however, depend on T only. For, assume that there is a second representation

$$T = \hat{P} x^\beta \hat{Q} = P x^\alpha Q.$$

Then, from the maximal choice of the α it follows

$$\beta_k \leq \alpha_k.$$

and from

$$\det T = \det \hat{Q}(0) x^{\sum \beta_k} + \dots = \det Q(0) x^{\sum \alpha_k} + \dots$$

one has

$$\sum_{k=1}^n \beta_k = \sum_{k=1}^n \alpha_k.$$

These two relations imply $\beta_k = \alpha_k$.

One can give an independent description for the α_k : Consider any subdeterminant Δ of order l of T and let

$$\Delta = c x^{\gamma_l} + \dots \quad c \neq 0$$

be the expansion of Δ . Then

$$\alpha_1 + \dots + \alpha_l = \text{Min } \gamma_l$$

where the minimum is taken over all l by l subdeterminants of T . This defines the α_k . The proof of this statement follows immediately from Lemma 1.

For any matrix $T(x)$ of the type occurring we define the span $\sigma(T)$ by

$$\sigma(T) = \alpha_n - \alpha_1 = \text{Max}_{k,l} (\alpha_k - \alpha_l)$$

which is a non negative integer. The assumption $\sigma(T) = 0$, for example, amounts to

$$T(x) = x^{\alpha_1} Q(x); \quad \det Q(0) \neq 0.$$

Any matrix T of span $\sigma > 0$ can be written as the product of σ matrices of span 1. For this purpose observe that x^α can be broken up into σ matrices of span 1 by constructing σ sequences $\beta^{(\nu)} = (\beta_k^{(\nu)})$ ($k = 1, \dots, n$; $\nu = 1, \dots, \sigma$) with

$$\beta_k^{(\nu)} \leq \beta_{k+1}^{(\nu)}; \quad \beta_n^{(\nu)} - \beta_1^{(\nu)} = 1.$$

and

$$\sum_{\nu=1}^{\sigma} \beta_k^{(\nu)} = \alpha_k.$$

Defining

$$T_\nu = P x^{\beta^{(\nu)}} P^{-1} \quad \text{for } \nu = 1, 2, \dots, \sigma - 1$$

$$T_\sigma = P x^{\beta^{(\sigma)}} Q$$

one has

$$\sigma(T^{(\nu)}) = 1.$$

$$T = T_1 T_2 \dots T_\sigma.$$

For later reference we introduce the ring $R[x]$ of functions $f(x)$ which are analytic in some (individual) neighborhood $|x| < \rho$. $R(x)$ denotes the corresponding field, generated by $R[x]$. The elements in $R(x)$ can be written in the form $x^\alpha f(x)$ where α is an integer and $f \in R[x]$, $f(0) \neq 0$. With $M[x]$, $M(x)$ we denote the ring of matrices whose elements lie in $R[x]$, $R(x)$ respectively. Thus $M[x]$, $M(x)$ form non commutative rings with zero divisors. The unit elements in the ring $M[x]$, consist of those matrices $Q(x)$ for which $\det Q(0) \neq 0$.

Similarly let $R_0[x]$ denote the ring of polynomials of x and $R_0(x)$ the field generated by $R_0[x]$. If $M_0[x]$, $M_0(x)$ are the rings of matrices whose elements are in $R_0[x]$, $R_0(x)$ one can express Lemma 1 as follows: Every $T(x) \in M(x)$ can be written in the form $P(x)x^\alpha Q(x)$ where P is a unit in $M_0[x]$ and Q is a unit in $M[x]$. x^α has the same meaning as above.

LEMMA 2. Let

$$A(x) = A_0 + xA_1 + \dots$$

belong to $M[x]$ and satisfy $r = r(A_0) > 0$ (i.e. $A_0 \neq 0$). Let $T(x)$ be a matrix in $M(x)$ with $\det T(x) \neq 0$ and define

$$B(x) = T^{-1}AT.$$

A necessary and sufficient condition for the existence of such a $T(x)$ that

$$B(x) = B_0 + xB_1 + \dots$$

belongs to $M[x]$ and satisfies

$$r(B_0) < r(A_0)$$

is that the polynomial

$$\mathfrak{P}(\lambda) = x^r \det \left(\lambda I + \frac{A(x)}{x} \right) \Big|_{x=0}$$

vanishes identically in λ .

PROOF. This condition is certainly necessary: If there exists such a $T(x)$ then

$$\begin{aligned} \det \left(\lambda I + \frac{A(x)}{x} \right) &= \det \left(\lambda I + \frac{B(x)}{x} \right) = \\ &= \det \left(\lambda I + \frac{B_0}{x} + B_1 + \dots \right). \end{aligned}$$

The last determinant contains x^{-1} at most to the power $r(B_0) < r$, hence

$$x^r \det \left(\lambda I + \frac{A(x)}{x} \right)$$

vanishes identically in λ for $x=0$.

To prove the converse we make use of the theory of invariants of matrices under similarity transformations ⁴⁾.

⁴⁾ The following proof is a slight modification of an argument for which I am indebted

to Professor A. A. ALBERT.

Let

$$f(\lambda, x) = \det(\lambda I + A(x))$$

be the characteristic polynomial of $A(x)$. Then the condition $\mathfrak{P}(\lambda) \equiv 0$ is tantamount to

$$x^{n-r+1} | f(\lambda x, x)$$

i.e. to the assumption that $x^{-n+r-1} f(\lambda x, x)$ is a power series in x . This follows immediately from the equation

$$x^{-n+r-1} f(\lambda x, x) = x^{r-1} \det\left(\lambda I + \frac{A(x)}{x}\right).$$

Denote by s the largest integer for which $x^s | f(\lambda x, x)$. Then $n - r < s \leq n$. We will prove the converse of Lemma 2 in the stronger version, that the rank of B_0 satisfies

$$(2.2^*) \quad r(B_0) \leq n - s;$$

which implies $r(B_0) < r(A_0)$.

To prove this statement we make use of the fact that a matrix $C = C(x)$ in $M(x)$ is similar to $A(x)$ if and only if it has the same elementary divisors $f_j(\lambda, x)$ as $A(x)$. The $f_j(\lambda, x)$ are polynomials in λ with coefficients in the field $R(x)$ which satisfy

$$(2.3) \quad f(\lambda, x) = \prod_{j=1}^{\omega} f_j(\lambda, x).$$

Since the highest coefficient of $f_j(\lambda, x)$ is one and concludes from Gauss' Lemma that the coefficients of $f_j(\lambda, x)$ lie in the ring $R[x]$.

Corresponding to the factorization of $f(\lambda, x)$ one can decompose the n dimensional vector space into n_j dimensional subspaces ($n = \sum_{j=1}^{\omega} n_j$) which are invariant under A . In the invariant subspaces A is given by matrices A_j with respect to an appropriate basis and

$$f_j(\lambda, x) = \det(\lambda I_{n_j} + A_j).$$

Let s_j be the largest integer satisfying $x^{s_j} | f_j(\lambda x, x)$ then it follows from (2.3) that $\sum_{j=1}^{\omega} s_j = s$.

Assuming that the statement (2.2*) has been proven for the indecomposable matrices A_j ⁵⁾, namely that they are similar to $B_j(x)$ in $M[x]$ with $r(B_j(0)) \leq n_j - s_j$ it follows that the matrix $B(x)$, which is composed of the $B_j(x)$, satisfies

$$r(B(0)) = \sum_{j=1}^{\omega} r(B_j(0)) \leq \sum_{j=1}^{\omega} (n_j - s_j) = n - s.$$

Therefore it suffices to prove the statement (2.2*) for indecomposable matrices $A(x)$. Then

$$(2.4) \quad f(\lambda, x) = \lambda^n + a_n(x) \lambda^{n-1} + \dots + a_1(x)$$

⁵⁾ For the concepts used here compare [5], Chap. IV.

is the minimal polynomial and the condition $x^s/f(\lambda x, x)$ ensures that

$$b_1(x) = \frac{a_1}{x^s}, \quad b_2 = \frac{a_2}{x^{s-1}}, \dots, b_s = \frac{a_s}{x}, \quad b_{s+1} = a_{s+1}, \dots, a_n = b_n.$$

are power series in $R[x]$. Any indecomposable matrix $A(x)$ satisfying (2.4) is similar to the so-called companion matrix $C(x)$ of $f(\lambda, x)$ which contains ones in the diagonal above the main diagonal and has as last row the elements

$$(a_1, a_2, \dots, a_n) = (b_1 x^s, b_2 x^{s-1}, \dots, b_s x, b_{s+1}, \dots, b_n).$$

All other elements of C are 0.

Defining the diagonal matrix

$$D(x) = (x^{-s}, x^{-s+1}, \dots, x^{-1}, 1, \dots, 1)$$

and

$$D^{-1}CD = B(x)$$

it is easily seen that $B(x)$ belongs to $M[x]$ and that the first s rows of $B(0)$ are zero rows, hence

$$r(B(0)) \leq n - s$$

which proves the statement (2.2*) and the Lemma 2.

LEMMA 3. Let $A(x)$ satisfy the assumption of Lemma 2 and assume

$$\mathfrak{P}(\lambda) \equiv 0.$$

Then there exists a transformation $T(x)$ (as in Lemma 2) for which

$$T^{-1}AT = B(x) \in M[x] \quad r(B(0)) < r(A(0))$$

for which

$$\sigma(T) = 1.$$

PROOF. a) For the proof it suffices to assume that $A(0)$ is nilpotent. Otherwise $A(0)$ has a nonzero eigenvalue $a \neq 0$ and with a constant similarity one can achieve that $A(0)$ is a triangular matrix and $a_{11} = a \neq 0$. Let \hat{A} be the $(n-1)$ by $(n-1)$ matrix which is obtained from $A(x)$ by cancelling the first row and column. It is easily seen that

$$\mathfrak{P}(\lambda) = x^r \det \left(\lambda I + \frac{A(x)}{x} \right) \Big|_{x=0} = a \hat{\mathfrak{P}}(\lambda)$$

where

$$\hat{\mathfrak{P}}(\lambda) = x^{r-1} \det \left(\lambda I + \frac{\hat{A}(x)}{x} \right) \Big|_{x=0}$$

and thus \hat{A} satisfies the same hypotheses as A does. If the statement can be proven for all matrices with $< n$ rows and columns it follows for n by n matrices provided there is a nonzero eigenvalue of $A(0)$. This reduces the proof to nilpotent $A(0)$.

b) By Lemma 2 there exists some $T(x) \in M(x)$ which diminishes the rank of the constant term. Representing $T(x)$ in the form

$$T(x) = P(x) x^\alpha Q(x)$$

according to Lemma 1 we define

$$\hat{A} = P^{-1} A P, \quad \hat{B} = Q B Q^{-1}$$

so that

$$x^{-\alpha} \hat{A} x^\alpha = \hat{B}.$$

Since

$$r(\hat{A}(0)) = r(A(0)); \quad r(\hat{B}(0)) = r(B(0))$$

we can assume $T(x) = x^\alpha$, where

$$\alpha = \text{diag}(\alpha_1, \dots, \alpha_n), \quad \alpha_1 \leq \alpha_2 \leq \dots \leq \alpha_n.$$

Furthermore, since $\mathfrak{B}(\lambda)$ depends on A_0, A_1 only we can assume $A_2 = A_3 = \dots = 0$.

c) In the decomposition $P x^\alpha Q$ the factors P, Q are not uniquely determined since there are polynomials P_1, P_2 for which

$$P_1 x^\alpha = x^\alpha P_2$$

for instance, for all P_1 which are upper triangular matrices $x^{-\alpha} P_1 x^\alpha$ are polynomials in x again.

This freedom in P, Q can be used to achieve that the nonzero column vectors of $A(0)$ are linearly independent. Thus if $A(0)$ is decomposed into several matrices with n rows:

$$A(0) = (\tilde{A}_1, \dots, \tilde{A}_r)$$

the rank is additive

$$r(A(0)) = \sum_{\nu=1}^r r(\tilde{A}_\nu).$$

To achieve this aim, we write

$$A(0) = (a_{kl})$$

and notice that

$$a_{kl} = 0 \quad \text{for } \alpha_k > \alpha_l,$$

since otherwise $x^{-\alpha} A x^\alpha$ would have a singularity at 0. We represent $A(0)$ by blocks of matrices by combining the a_{kl} with $\alpha_k = \alpha, \alpha_l = \alpha'$ to $A_{\alpha\alpha'}$ so that

$$A(0) = (A_{\alpha\alpha'})$$

where α, α' runs over certain integers. The above condition amounts to

$$A_{\alpha\alpha'} = 0 \quad \text{for } \alpha > \alpha'.$$

The diagonal matrices $A_{\alpha\alpha}$, of course, are nilpotent since $A(0)$ is. By a constant transformation

$$T_0 = (T_{0\alpha\alpha'})$$

which is a diagonal matrix in this block representation ($T_{\alpha\alpha'} = 0$ for $\alpha \neq \alpha'$) one can achieve that $A_{\alpha\alpha}$ are in upper triangular normal form. Thus $A(0) = (a_{ki})$ are upper triangular with zero diagonal elements.

Now we achieve that the column vectors of $A(0)$ are replaced by 0 vectors and r independent vectors. For this purpose let a_i denote the column vectors of $A(0)$;

$$A(0) = (a_1, a_2, \dots, a_n).$$

Let a_1, a_2, \dots, a_{l-1} consist of zero and independent vectors already and assume a_l is dependent on a_1, a_2, \dots, a_{l-1}

$$a_l = \sum_{\lambda < l} c_\lambda a_\lambda.$$

Let C denote the constant matrix which is obtained from the unit matrix by replacing the l^{th} column by

$$\begin{pmatrix} -c_1 \\ -c_2 \\ \vdots \\ -c_{l-1} \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

Then

$$A(0)C = (a_1, a_2, \dots, a_{l-1}, 0, a_{l+1}, \dots, a_n).$$

Since $C^{-1}A(0)C$ is obtained from $A(0)C$ by adding the l^{th} row to the λ^{th} ($\lambda < l$) the first l columns are not affected by this procedure since $A(0)C$ like $A(0)$ is upper triangular with zero diagonal elements. Also since

$$x^{-\alpha} C x^\alpha = C_1(x)$$

is a polynomial with $\det C_1(x) \equiv 1$ (hence a unit in the polynomial ring) we have

$$x^{-\alpha} (C^{-1} A C) x^\alpha = C_1^{-1} B C_1 = \hat{B}$$

$$r(C^{-1} A(0) C) = r(A(0))$$

$$r(\hat{B}(0)) = r(B(0)).$$

Thus we can assume that A is replaced by $C^{-1}AC$ in the above statement.

Thus we can assume that the rank is additive with respect to a decomposition of $A(0)$ into several column matrices.

d) For the sake of simplicity we first carry out the proof of Lemma 3 in case

$$\alpha_1 < \alpha_2 < \dots < \alpha_n.$$

The general case will be a generalization of this idea.

By a), b), c) we can assume

$$A = A_0 + x A_1$$

A_0 consists of 0 and independent column vectors

A_0 is upper triangular.

Since

$$B = x^{-\alpha} A x^\alpha$$

$$B_{kl} = x^{-\alpha_k + \alpha_l} A_{kl}(x)$$

the elements of $B(0)$ are zero except if $-\alpha_k + \alpha_l = -1$. Thus with

$$B(0) = (b_{kl})$$

one has

$$b_{kl} = 0 \quad \text{for } \alpha_k \neq \alpha_l + 1$$

in particular,

$$b_{kl} = 0 \quad \text{for } k \neq l + 1.$$

The α_k form an increasing sequence of integers. We construct all those sequences $\beta_k^{(\nu)}$ for which $\beta_n^{(\nu)} - \beta_1^{(\nu)} = 1$ which have the jump from $k = \nu - 1$ to $k = \nu$ ($\nu = 2, \dots, n$). Introducing

$$\beta^{(\nu)} = \text{diag}(\beta_1^{(\nu)}, \beta_2^{(\nu)}, \dots, \beta_n^{(\nu)}) \quad (\nu = 2, \dots, n)$$

we form the matrices

$$x^{-\beta^{(\nu)}} A(x) x^{\beta^{(\nu)}} = C^{(\nu)}(x)$$

and claim that

$$(*) \quad r(C^{(\nu)}(0) < r(A_0))$$

for some ν . Counterexamples show that in general this inequality is not true for all ν .

To prove this statement notice that $\beta_k^{(\nu)} - \beta_l^{(\nu)} > 0$ implies $\alpha_k - \alpha_l > 0$. Therefore $C^{(\nu)}(x)$ has no negative powers of x and the matrix $C^{(\nu)}(0)$ has the form

$0 \ a_{12} \ \dots \ a_{1, \nu-1}$ \vdots $\dots \ a_{\nu-2, \nu-1}$	0
$0 \ \dots \ 0$	0
$b_{\nu, \nu-1}$	$0 \ a_{\nu, \nu+1}$ \dots $a_{n-1, n}$ 0

Hence the rank of $C^{(\nu)}(0)$ is at most

$$r(a_1, a_2, \dots, a_{\nu-1}, a_{\nu+1}, \dots, a_n) + r(b_\nu)$$

where b_ν stands for the ν^{th} row of $B(0)$. Since the rank is additive for A_0 it follows

$$r(C^{(\nu)}) \leq r(A_0) - r(a_\nu) + r(b_\nu).$$

Adding over all ν one obtains

$$\sum_{\nu=2}^n r(C^{(\nu)}) \leq (n-1)r(A_0) - r(A_0) + r(B(0)).$$

The assumption $r(A_0) > r(B(0))$ leads to

$$\frac{1}{n-1} \sum_{\nu=2}^n r(C^{(\nu)}) < r(A_0)$$

which proves the statement (*).

e) A similar argument as in d) can be applied in the general case

$$\alpha_1 \leq \alpha_2 \leq \dots \leq \alpha_n.$$

In this case, however, one first has to achieve that $B(0)$ has the additive rank property with respect to the rows and simultaneously $A(0)$ with respect to the columns. Here we refer to rows and columns in the block representation determined by indices k with equal α_k . In other words: Let

$$A_0 = (A^{(1)}, A^{(2)}, \dots, A^{(N)})$$

be a decomposition of A_0 into N matrices of consecutive columns in such a manner that $\alpha_k \neq \alpha_{k'}$ if k, k' are the column numbers belonging to different matrices $A^{(K)}$. Then

$$r(A_0) = \sum_{K=1}^N r(A^{(K)}).$$

A similar property can be achieved simultaneously with respect to the rows of $B(0)$.

To prove this statement we can assume $A(0)$ to be an upper triangular matrix consisting of r independent column vectors and $n-r$ zero columns. Writing $A(0)$ in the block representation

$$A(0) = (A_{\alpha\alpha'}), \quad B(0) = (B_{\alpha\alpha'})$$

where α, α' runs over the integers which are assumed by $\alpha_1, \dots, \alpha_n$ one has in particular, $A_{\alpha\alpha'} = 0$ for $\alpha > \alpha'$. Breaking up $A(0)$ into column matrices

$$A_{\alpha'} = (A_{\alpha\alpha'})$$

one obviously has

$$(**) \quad r(A(0)) = \sum_{\alpha'} r(A_{\alpha'}).$$

In order to achieve the corresponding property with respect to the rows of $B(0)$ we apply a constant matrix

$$C = (C_{\alpha\alpha'}) \quad (C_{\alpha\alpha'} = 0 \text{ if } \alpha \neq \alpha')$$

(diagonal in block representation) so that $A_{\alpha\alpha} = B_{\alpha\alpha}$ and hence $B(0)$ become lower triangular matrices. Since C commutes with x^α one has

$$x^{-\alpha}(C^{-1}AC)x^\alpha = C^{-1}BC$$

and the matrix $C^{-1}A(0)C$ satisfies (**) since the rank of A_α remains unchanged.

Applying the construction of c) we achieve that $B(0)$ consists of zero and linear independent row vectors. For this purpose we construct a constant matrix C_1 such that

$$C_1 B(0) C_1^{-1}$$

consists of zero rows and linearly independent rows. C_1 has ones in the diagonal and zeros above the diagonal. Then $x^{-\alpha}A x^\alpha = B$ goes over into

$$x^{-\alpha}C_2 A C_2^{-1} x^\alpha = C_1 B C_1^{-1}$$

where

$$C_2(x) = x^\alpha C_1 x^{-\alpha}$$

$C_2(x)$ is regular at $x=0$, since C_1 has zeros above the diagonal. Moreover $C_2(0)$ is nonsingular and $\det C_2(0) = 1$. Actually $C_2(0)$ is diagonal in the block representation so that the condition (**) will be satisfied for $C_2 A C_2^{-1}$ as for A . This shows that one can assume that A as well as B have the additive rank property with respect to their columns (A) and their rows (B) respectively.

f) Now the proof can be finished like in d). Let $\beta^{(\nu)} = \text{diag}(\beta_1^{(\nu)}, \beta_2^{(\nu)}, \dots, \beta_n^{(\nu)})$ ($\nu = 2, \dots, N$), where $\beta_k^{(\nu)}$ are increasing sequences of 0 and 1 for which the jump occurs only at a jump of the given sequence α_k . Then $T = x^{\beta^{(\nu)}}$ have span 1. Introducing

$$x^{-\beta^{(\nu)}} A(x) x^{\beta^{(\nu)}} = C^{(\nu)}(x) \quad (\nu = 2, \dots, N)$$

it can be shown that

$$r(C^{(\nu)}(0)) = r(A_0) - r(A^{(\nu)}) + r(B^{(\nu)})$$

where

$$A_0 = (A^{(1)}, A^{(2)}, \dots, A^{(N)}), \quad B_0 = \begin{pmatrix} B^{(1)} \\ \vdots \\ B^{(N)} \end{pmatrix}.$$

Adding these relations one finds

$$\frac{1}{N-1} \sum_{\nu=2}^N r(C^{(\nu)}(0)) < r(A_0)$$

hence

$$r(C^{(\nu)}(0)) < r(A_0)$$

for at least one ν . Since $x^{\beta^{(\nu)}}$ has span 1 the Lemma 3 is proven.

III. Proof of Theorem 1 and 2

a) Assume $A = x^{-p}(A_0 + xA_1 + \dots)$ satisfies

$$\mathfrak{P}(\lambda) = x^r \det \left(\frac{A_0}{x} + A_1 + \lambda I \right) \Big|_{x=0} \equiv 0.$$

Then according to Lemma 3 there is a transformation

$$T(x) = T_0 + xT_1 + \dots = T$$

with $\sigma(T) = 1$ such that

$$T^{-1}AT = B$$

and

$$m(A) > m(B).$$

Since $\sigma(T) = 1$ one can write T in the form $Px^\alpha Q$ with $\alpha_n - \alpha_1 = 1$. Hence

$$T^{-1} \frac{d}{dx} T = Q^{-1} x^{-\alpha} P^{-1} \frac{dP}{dx} x^\alpha Q + Q^{-1} \frac{\alpha}{x} Q + Q^{-1} \frac{dQ}{dx}$$

has at most a pole of order 1. Thus, if $p > 1$ (i.e. $m > 1$) then

$$m \left(T^{-1}AT - T^{-1} \frac{d}{dx} T \right) < m(A)$$

which proves the first part of the theorem.

b) If on the other hand there exists some $T(x)$ such that

$$T^{-1}AT - T^{-1} \frac{d}{dx} T = B$$

and

$$m(B) < m(A)$$

then one can represent T in the form

$$Px^\alpha Q.$$

Then with

$$P^{-1}AP - P^{-1} \frac{d}{dx} P = \hat{A}$$

$$QBQ^{-1} + \left(\frac{d}{dx} Q \right) Q^{-1} = \hat{B}$$

one has

$$x^{-\alpha} \hat{A} x^\alpha - \frac{\alpha}{x} = \hat{B}$$

and

$$m(\hat{A}) = m(A); \quad m(\hat{B}) = m(B)$$

since P, Q are regular at $x=0$. Since we assumed $m(A) > 1$ it follows

$$m(\hat{A}) > m \left(\hat{B} + \frac{\alpha}{x} \right)$$

thus m can be diminished by a similarity transformation. This proves that

$$x^r \det (\lambda I + x^{(p-1)} \hat{A}) \Big|_{x=0} \equiv 0.$$

Since this expression depends on A_0, A_1 only and since $\hat{A} - P^{-1}AP = -P^{-1} \frac{d}{dx} P$ is a power series in x it follows that $\mathfrak{P}(\lambda) = x^r \det(\lambda I + x^{p-1}A)|_{x=0} \equiv 0$. In other words: If the number $m(A)$ is greater than 1 it can be diminished by similarity if and only if it can be diminished by the operation

$$T^{-1}AT - T^{-1}T'$$

The transformation T can always be chosen as to be of span 1. Moreover, representing T in the form $Px^\alpha Q$ of Lemma 1 one observes that $Q(x)$ does not change $m(B)$ and therefore can be replaced by the identity. Finally it is obvious that only the first two terms of $P(x)$ enter in B_0 , so that we can assume $T = (P_0 + P_1x)x^\alpha$. This completes the proof of theorem 2.

In particular one notices that for symmetric matrices $A(x)$ only the reduced form can occur. Since for a symmetric matrix the rank agrees with the number of non zero eigenvalues one has

$$\mathfrak{P}(\lambda) = x^r \det\left(\lambda I + \frac{A_0}{x} + A_1\right)\Big|_{x=0} = a_1, \dots, a_r \det(\lambda I + \hat{A})$$

if a_1, \dots, a_r are the non zero eigenvalues of A_0 . This expression is a polynomial of degree $n - r$.

IV. The case of a single differential equation

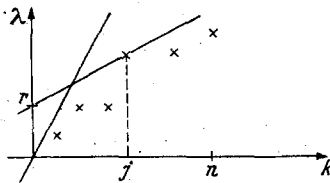
Consider a differential equation of n -th order

$$(E) \quad \frac{d^n u}{dx^n} + a_1(x) \frac{d^{n-1} u}{dx^{n-1}} + \dots + a_{n-1}(x) \frac{du}{dx} + a_n(x) u = 0$$

for which the coefficients $a_k(x)$ have at most a pole at zero. Let $\lambda_k \geq 0$ denote the order of the pole of a_k . ($\lambda_k = 0$ if a_k is regular).

We now construct p, r as the smallest integers (p, r) ordered lexicographically such that

$$\lambda_k \leq k(p - 1) + r.$$



This can be visualized graphically if one plots the λ_k as ordinates over the abscissa k and then first constructs the straight line through 0 with the smallest integer slope p which stays above all λ_k . Then one finds among all the parallel lines of slope $p - 1$ the lowest one. The section which is cut off on the λ -axis is r , which is an integer. That r is an integer follows from the formula

$$r = \text{Max}_k (\lambda_k - (p - 1)k).$$

Defining

$$\gamma_k = \text{Min}(pk, (p - 1)k + r)$$

we have

$$\lambda_k \leq \gamma_k$$

and equality for at least one $k \geq r$. Let $j \geq r$ be the smallest integer with

$$\lambda_j = \gamma_j.$$

Geometrically γ_k represents the broken line dominating λ_k .

Now it is easy to construct an irreducible system of differential equations which is equivalent to (E). For this purpose introduce the vector y with the components

$$y_{k+1} = x^{-\gamma_{n-k}} \frac{d^k u}{dx^k} \quad (k = 0, \dots, n-1).$$

Then for $k = 0, \dots, n-2$

$$\begin{aligned} \frac{dy_{k+1}}{dx} &= x^{-\gamma_{n-k} + \gamma_{n-k-1}} y_{k+2} - \frac{\gamma_{n-k}}{x} y_{k+1} \\ \frac{dy_n}{dx} &= - \sum_{k=0}^{n-1} a_{n-k} x^{-\gamma_1 + \gamma_{n-k}} y_{k+1} - \frac{\gamma_1}{x} y_n. \end{aligned}$$

Therefore, introducing

$$\gamma = \text{diag}(\gamma_n, \dots, \gamma_1)$$

and

$$K(x) = \left(\begin{array}{ccc|cc} \overbrace{}^{n-r+1} & & & & \\ 0 & x & & & \\ & & \ddots & & \\ & & & x & \\ & & & 0 & 1 \\ & & & & 0 \\ & & & & & 1 \\ b_n, b_j \dots b_r & & & b_{r-1} & b_1 \end{array} \right)$$

where the b_k are the regular functions

$$b_k(x) = -a_k(x) x^{\gamma_k} \quad (k = 1, \dots, n)$$

we have

$$\frac{dy}{dx} = \left(x^{-p} K(x) - \frac{\gamma}{x} \right) y^{\text{e}}.$$

It has to be shown that for

$$A(x) = x^{-p} K(x) - \frac{\gamma}{x}$$

the polynomial $\mathfrak{P}(\lambda)$ does not vanish identically, if $p > 1$. It is obvious that

$$x^p A(x)|_{x=0} = A_0 = K(0)$$

has rank r , since $r-1$ rows of $K(0)$ contains ones and since

$$b_j(x) = -a_j(x) x^{\gamma_j} = -a_j(x) x^{\lambda_j} \quad (j \geq r)$$

^e) Use $\gamma_1 = \text{Min} \{p, p-1+r\} = p$.

does not vanish at $x=0$. If one considers the determinant

$$\mathfrak{P}(\lambda) = x^r \det \left(\lambda I + \frac{1}{x} K(x) - \gamma x^{p-2} \right) \Big|_{x=0}$$

one computes easily the highest coefficient to be

$$\mathfrak{P}(\lambda) = \lambda^{n-i} b_j(0) + \dots$$

where $j \geq r$ was defined above. Hence the degree of $\mathfrak{P}(\lambda)$ is $\leq n-r$ and $\mathfrak{P}(\lambda) \neq 0$. This proves that the above system is not reducible, hence μ is given by

$$\mu = p - 1 + \frac{r}{n} \quad \text{if } p > 1$$

where p, r were defined by the above diagram.

Necessary and sufficient for a regular singularity is that $p=1$ or $\mu \leq 1$ which is the principal result of Fuchs theory. The above result can be considered a generalization of Fuchs' theorem.

References

- [1] HORN, J.: Zur Theorie der Systeme linearer Differentialgleichungen mit einer unabhängigen Veränderlichen, II. *Math. Annalen* **40**, 527–550 (1892). — [2] BIRKHOFF, G. D.: Equivalent singularities of ordinary linear differential equations. *Math. Ann.* **74**, 134–139 (1913). — [3] BREUERBACH, L.: Theorie der gewöhnlichen Differentialgleichungen. Berlin: Springer 1953; esp. § 6. — [4] CODDINGTON, E. A., and N. LEVINSON: Theory of Ordinary Differential Equations. McGraw-Hill 1955; Chap. 4 and 5. — [5] ALBERT, A. A.: Modern Algebra. Chicago 1937, 1958; in particular Chap. IV. — [6] LOEWY, A.: Über Matrizen- und Differentialkomplexe, I, II, III. *Math. Ann.* **78**, 1–51, 343–358, 359–369 (1918). — [7] PERRON, O.: Über diejenigen Integrale linearer Differentialgleichungen, welche sich an einer Unbestimmtheitsstelle bestimmt verhalten. *Math. Ann.* **70**, 1–32 (1911).

Dept. of Math., Massachusetts Institute of Technology, Cambridge 39, Mass. (U.S.A.)

(Eingegangen am 2. Mai 1959)