YOKOHAMA MATHEMATICAL
JOURNAL VOL. 29, 1981

THE DETAILED STRUCTURE OF FORMAL ANALYTIC INVARIANTS OF A MEROMORPHIC DIFFERENTIAL EQUATION

By

W. BALSER

(Received June 16, 1980. Revised November 28, 1980)

0. Introduction

In [1, 2], [4] complete systems of invariants corresponding to several different types of equivalence for meromorphic differential equations $x'=A(z)x$ have been defined, and the author(s) have been aiming at making the structure of these invariants as explicit as possible. For example, the system of invariants corresponding to *formal meromorphic equivalence* can be thought of as a matrix $G_m(z)$ (the *formal meromorphic invariant*) which is uniquely defined and explicitly calculable in terms of $A(z)$. Furthermore, distinguishing several different cases in which $G_m(z)$ shows a different structure, one may see that in everyone of these cases $G_m(z)$ depends upon finitely many parameters which vary independently within their range of definition and whose number can be found explicitly in each case.

For example, if n (the size of the coefficient matrix $A(z)$) equals 3 and the Poincaré rank of $A(z)$ is one, a complete description of $G_m(z)$ can be given as follows:

$$G_m(z)=z^J U \exp [Q(z)],$$

where

a) $Q(z)=\operatorname{diag}[\lambda_1 z, \lambda_2 z, \lambda_3 z]$,
$\operatorname{Re}\lambda_j<\operatorname{Re}\lambda_k$ or $(\operatorname{Re}\lambda_j=\operatorname{Re}\lambda_k, \operatorname{Im}\lambda_j<\operatorname{Im}\lambda_k)$ for $1\leqq j<k\leqq 3$,
$U=I$, $J=\operatorname{diag}[\lambda_1', \lambda_2', \lambda_3']$, $0\leqq\operatorname{Re}\lambda_j'<1$ $(1\leqq j\leqq 3)$.

b) $Q(z)=\operatorname{diag}[\lambda_1 z, \lambda_2 z+\tilde{\lambda}z^{1/2}, \lambda_2 z-\tilde{\lambda}z^{1/2}]$, $\lambda_1\neq\lambda_2$, $\operatorname{Re}\tilde{\lambda}>0$,

$$U=\begin{bmatrix}1 & 0 & 0\\ 0 & 1 & 1\\ 0 & 1 & -1\end{bmatrix},$$

$$J=\operatorname{diag}\left[\lambda_1', \lambda_2', \lambda_2'+\frac{1}{2}\right],\quad 0\leqq\operatorname{Re}\lambda_1'<1,\quad 0\leqq\operatorname{Re}\lambda_2'<\frac{1}{2}.$$

c) $Q(z)=\operatorname{diag}[\lambda_1 z, \lambda_2 z, \lambda_2 z]$, $\lambda_1\neq\lambda_2$,

$U=I, J$ as in a) with $\mathrm{Re}\,\lambda_2'<\mathrm{Re}\,\lambda_3'$ or ($\mathrm{Re}\,\lambda_2'=\mathrm{Re}\,\lambda_3'$, $\mathrm{Im}\,\lambda_2'\leqq\mathrm{Im}\,\lambda_3'$).

d) $Q(z)$, U as in c),

$$J=\begin{bmatrix}\lambda_1' & 0 & 0\\ 0 & \lambda_2' & 0\\ 0 & 1 & \lambda_2'\end{bmatrix},\quad 0\leqq \mathrm{Re}\,\lambda_j'<1\ (j=1,2).$$

e) $Q(z)=\mathrm{diag}\,[q(z), q(z\varepsilon), q(z\varepsilon^2)]$,
$q(z)=\lambda^{(0)}z+\lambda^{(1)}z^{2/3}+\lambda^{(2)}z^{1/3}$, $\varepsilon=e^{2\pi i/3}$,
either $\lambda^{(1)}\neq 0$, $0\leqq\arg\lambda^{(1)}<2\pi/3$,
or $\lambda^{(1)}=0$, $\lambda^{(2)}\neq 0$, $0\leqq\arg\lambda^{(2)}<2\pi/3$,

$$U=\begin{bmatrix}1 & 1 & 1\\ 1 & \varepsilon & \varepsilon^2\\ 1 & \varepsilon^2 & \varepsilon^4\end{bmatrix},\quad J=\mathrm{diag}\left[\lambda', \lambda'+\frac{1}{3}, \lambda'+\frac{2}{3}\right],\quad 0\leqq\mathrm{Re}\,\lambda'<\frac{1}{3}.$$

f) $Q(z)=\lambda zI$, $U=I$,
J as in a) with $\mathrm{Re}\,\lambda_j'<\mathrm{Re}\,\lambda_k'$ or ($\mathrm{Re}\,\lambda_j'=\mathrm{Re}\,\lambda_k'$, $\mathrm{Im}\,\lambda_j'\leqq\mathrm{Im}\,\lambda_k'$) $(1\leqq j<k\leqq 3)$.

g) $Q(z)$, U as in f), J as in d).

h) $Q(z)$, U as in f),

$$J=\begin{bmatrix}\lambda' & 0 & 0\\ 1 & \lambda' & 0\\ 0 & 1 & \lambda'\end{bmatrix},\quad 0\leqq\mathrm{Re}\,\lambda'<1\,.$$

In an analogous way, one can write out the different cases occuring for $G_m(z)$ for larger n (and higher Poincaré rank), although the number of cases increases rather rapidly with n, and in every case the formal meromorphic invariant has a completly explicit structure.

The situation in case of *formal analytic equivalence* is much less explicit: A formal analytic invariant corresponding to $x'=A(z)x$ has been defined in [1], [4] as a matrix of the form

$$G_a(z)=P(z)z^K G_m(z)\,,$$

where K is a diagonal matrix of integer entries, and $P(z)$ is lower triangular with ones along the diagonal and polynomials in z without constant term below the diagonal. In order to define a *unique* $G_a(z)$ corresponding to $x'=A(z)x$, a representative was a-priori selected out of the set of all possible $P(z)z^K$ occuring with a given differential equation. This set turned out to be an equivalence class corresponding to the following equivalence definition:

A matrix $\tilde{P}(z)z^{\tilde{K}}$ of the same type as $P(z)z^K$ is said to be equivalent to $P(z)z^K$ (relative to $G_m(z)$) whenever a constant invertible C commuting with $G_m(z)$ and an analytic transformation $C(z)$ exist such that

$$P(z)z^{K}C=C(z)\tilde{P}(z)z^{\tilde{K}}\,. \tag{0.1}$$

In contrast to the formal meromorphic case, the selection of a representative $P(z)z^{K}$ was done just by help of the axiom of choice. So for the formal analytic invariant one cannot directly determine the number of parameters contained in $G_a(z)$ (corresponding to differential equations of fixed size and Poincaré rank, say).

In [3], the author contributed an algorithmic procedure (in case $n=2$) which, given any $P(z)z^{K}$ and the group $\mathscr{G}_m$ of constant invertible matrices commuting with G_m, enables to calculate a representative within the equivalence class of $P(z)z^{K}$: In case $n=2$, the group $\mathscr{G}_m$ is either

$$\mathscr{G}_m=\left\{\begin{bmatrix} c_1 & 0 \\ c_2 & c_1\end{bmatrix},\ c_1\neq 0\right\}, \tag{α}$$

or

$$\mathscr{G}_m=\left\{\begin{bmatrix} c_1 & 0 \\ 0 & c_2\end{bmatrix},\ c_1c_2\neq 0\right\}, \tag{β}$$

or

$$\mathscr{G}_m=\left\{\begin{bmatrix} c_{11} & c_{12} \\ c_{21} & c_{22}\end{bmatrix},\ c_{11}c_{22}-c_{12}c_{21}\neq 0\right\}. \tag{γ}$$

Then in every equivalence class there is a unique representative $P(z)z^{K}$,

$$P(z)=\begin{bmatrix} 1 & 0 \\ p(z) & 1\end{bmatrix}, \qquad K=\operatorname{diag}[l, k],\ l,\ k \text{ integer},$$

satisfying the following conditions:

In case (α): The polynomial $p(z)$ does not contain a term z^{k-l} (if $k\leqq l$, this is to be interpreted void, since by definition $p(z)$ is a polynomial without constant term).

In case (β): Either $p(z)\equiv 0$ or $p(z)$ has highest coefficient 1.

In case (γ): Either $p(z)\equiv 0$ and $k\leqq l$, or $k-l<\deg p$ and $p(z)$ has highest coefficient 1 and does not contain a term z^{k-l}.

In all three cases, if a given $P(z)z^{K}$ does not satisfy the conditions stated, then one can explicitly find $C\in\mathscr{G}_m$ such that $\tilde{P}(z)z^{\tilde{K}}$ defined by (0.1) satisfies the conditions and therefore $\tilde{P}(z)z^{\tilde{K}}$ is the unique representative.

The purpose of this paper is to generalize this result to arbitrary n. This is again done by stating an algorithmic procedure involving a finite number of steps of the following type.

Given any group $\mathscr{G}$ of invertible, constant matrices showing a structure described in *Section 1*, we consider an arbitrary equivalence class of matrices $P(z)z^K$ (relative to $\mathscr{G}$). If the equivalence class contains at least two different matrices, we specify a subset of matrices $P(z)z^K$ by fixing some of the parameters in $P(z)z^K$ to certain "natural" values, such as minimal values for the integers in K or zero or one for the coefficients of the polynomials in $P(z)$. It is a non-trivial consequence of the special structure of $\mathscr{G}$ that this selection of a subset can always be done in a way that the subset of matrices $P(z)z^K$ turns out to be an equivalence class relative to a subgroup of $\mathscr{G}$ which is essentially of the same structure as $\mathscr{G}$.

The following example shows some of the typical arguments which are later used in general:

Let G_m be as in case d) above (with $\lambda_1' \neq \lambda_2'$), and let $\mathscr{G}_m = \mathscr{G}$ be the group of constant invertible C commuting with G_m, i.e. C is of the form

$$C = \begin{bmatrix} c_1 & 0 & 0 \\ 0 & c_2 & 0 \\ 0 & c_3 & c_2 \end{bmatrix}, \qquad c_1 c_2 \neq 0 .$$

Let

$$P(z) = \begin{bmatrix} 1 & 0 & 0 \\ p_{21}(z) & 1 & 0 \\ p_{31}(z) & p_{32}(z) & 1 \end{bmatrix}, \qquad K = \operatorname{diag}[k_1, k_2, k_3],$$

then every $\tilde{P}(z)z^{\tilde{K}}$ equivalent to $P(z)z^K$ can be found by factorizing $P(z)z^K C$ (with C as above) into $C(z)\tilde{P}(z)z^{\tilde{K}}$ (with an analytic transformation $C(z)$), using a Gaussian algorithm type procedure (compare [4]). Since C is lower triangular, we find $\tilde{K} = K$, $\tilde{p}_{21}(z) = p_{21}(z)c_1/c_2$, $\tilde{p}_{31}(z) = p_{31}(z)c_1/c_2$, and $\tilde{p}_{32}(z)$ is the polynomial part (without constant term) of $p_{32}(z) + z^{k_3-k_2}c_3/c_2$. Hence assuming $p_{21} \neq 0$, as a first step we may consider the subset of matrices (of the given equivalence class) for which the highest coefficient of $p_{21}(z)$ is one. If both $P(z)z^K$ and $\tilde{P}(z)z^{\tilde{K}}$ are taken from this subset, then (0.1) holds with a C for which $c_1 = c_2$, and vice versa if $c_1 = c_2$ and $p_{21}(z)$ has highest coefficient one, then $\tilde{p}_{21}(z)$ has also highest coefficient equal to one. The set of $C \in \mathscr{G}$ with $c_1 = c_2$ forms a subgroup $\tilde{\mathscr{G}}$, and we next may consider which influence we have upon $p_{32}(z)$, when working with matrices $C \in \tilde{\mathscr{G}}$. We immediately see that $\tilde{p}_{32}(z) = p_{32}(z)$ if $k_3 - k_2 \leqq 0$, and $\tilde{p}_{32}(z) = p_{32}(z) + z^{k_3-k_2}c_3/c_2$ otherwise.

If the first case occurs, then the equivalence class (with respect to $\tilde{\mathscr{G}}$) contains only one element $P(z)z^K$ which we take as the representative, whereas in the second case we restrict to the subset of matrices where $p_{32}(z)$ contains no term with exponent $k_3 - k_2$. This subset, however, is an equivalence class relative to the subgroup of matrices with $c_3=0$ (and $c_1=c_2$), i.e. the subgroup of scalar matrices c_1I, and the corresponding equivalence class again contains only one element. Similar arguments apply if $p_{21}=0$.

In order to make analogous arguments work in general, it is most important to understand the structure of the group $\mathscr{G}_m$ and its subgroups occuring in the procedure. It was shown in [1], [4] that C commutes with $G_m(z)=z^J U \exp[Q(z)]$ iff it commutes with $Q(z)$, with a block permutation matrix R explicitly given in terms of $Q(z)$, and with J separately. Since J was taken to be a matrix in Jordan canonical form, the structure of $\mathscr{G}_m$ is related to the structure of matrices commuting with Jordan matrices. Although those matrices can be completely characterized ([4], pp. 49, 50), they do not have a structure which is convenient for handling our normalization problem. We therefore define a *modified canonical form* J, and show that the group of matrices commuting with J consists of all lower triangularly blocked matrices C (corresponding to a block structure determined by J) which satisfy additional restrictions expressed in terms of the blocks rather than of the single elements of C (Section 1). For the definition of the formal meromorphic invariant $G_m(z)$ it was only essential to fix a unique normal form corresponding to every constant matrix. This was done in [1], [4] by choosing Jordan canonical form together with an a-priori-ordering of the eigenvalues and the Jordan blocks corresponding to the same eigenvalue ([1], Section 1; [4], § 3e). Instead one might as well use the modified canonical form together with an a-priori-ordering of the eigenvalues (compare Remark 1.4), and the reader may check that none of the results concerning $G_m(z)$ fails to hold (modulo changes in notation) when replacing Jordan canonical forms by modified canonical forms.

In *Section 2* we (analogously to (0.1)) define the notation of $\mathscr{G}$-equivalence of matrices $P(z)z^K$ corresponding to an *admissible* group $\mathscr{G}$ (i.e. a group of matrices having a structure similar to those commuting with a matrix in modified canonical form, and we indicate several ways of defining subgroups of $\mathscr{G}$ which again are admissible.

Section 3 contains a discussion of what can be considered as a first step towards selecting a representative out of any $\mathscr{G}$-equivalence class of matrices $P(z)z^K$ in case $\mathscr{G}$ consists of all $n\times n$ constant, invertible matrices. In *Section 4* we then return to the general case and explain how one can select a unique representative $P(z)z^K$. In a final *Section* 5, we consider as an example the case $n=3$, and apply our results to the case of normal forms of constant matrices with respect to a restricted type of similarity.

We mention the following improvement in the theory of invariants obtained from this paper: Assume that any two meromorphic differential equations $X'=A(z)X$ and $\tilde{X}'=\tilde{A}(z)\tilde{X}$ are given, which shall be shown to be formally analytically equivalent (resp. inequivalent). Then by algebraic arguments one can find formal fundamental solutions $H(z)$ ($\tilde{H}(z)$) of either one of the equations, which are of the form

$$H(z)=F_a(z)P(z)z^K G_m(z)\ ,$$
$$(\text{resp. } \tilde{H}(z)=\tilde{F}_a(z)\tilde{P}(z)z^{\tilde{K}}\tilde{G}_m(z))\ ,$$

where $G_m(z)$ ($\tilde{G}_m(z)$) is the formal meromorphic invariant, $P(z)z^K(\tilde{P}(z)z^{\tilde{K}})$ is as described before, and $F_a(z)$ ($\tilde{F}_a(z)$) is a formal analytic transformation. Then it is necessary and sufficient for formal analytic equivalence of the two differential equations that $G_m(z)=\tilde{G}_m(z)$, and that (0.1) holds for a suitable constant, invertible matrix $C\in\mathscr{G}_m$. While the first condition is immediate to check, the second one will involve finding (resp. disproving the existence of) the matrix C. Applying the results of this paper, however, enables to find the representative corresponding to $P(z)z^K(\tilde{P}(z)z^{\tilde{K}})$ within finitely many straight forward steps, and then the question of whether (0.1) holds or not is decided by whether or not the two representatives are equal.

1. A modified canonical form

Corresponding to an arbitary constant matrix C, we define a *modified canonical form* J, which differs from the usual Jordan canonical form by a permutation similarity. We do this since matrices commuting with J (in modified canonical form) have a simpler structure. As an example, one may see from the following Lemma 1 that the matrix whose Jordan canonical form is

$$N=\begin{bmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

has the modified canonical form

$$J=\begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix},$$

and the matrices C commuting with J are precisely of the (blocked) form

$$C=\begin{bmatrix} C_1 & 0 \\ C_2 & C_1 \end{bmatrix}, \qquad C_1 \text{ of size } 2\times 2,$$

whereas those commuting with N have the form

$$C=\begin{bmatrix} c_1 & 0 & c_3 & 0 \\ c_2 & c_1 & c_4 & c_3 \\ c_5 & 0 & c_7 & 0 \\ c_6 & c_5 & c_8 & c_7 \end{bmatrix}, \qquad c_1, \cdots, c_8 \text{ arbitrary.}$$

Throughout this paper, I_s (for any natural s) denotes the identity matrix of size $s\times s$.

Lemma 1. *Let s be some natural number, then every constant $s\times s$ matrix C is similar to a matrix J having the following properties: If $\lambda_1, \cdots, \lambda_\mu$ denote the (distinct) eigenvalues of C, and $s_1, \cdots, s_\mu$ the multiplicities of the corresponding eigenvalues, then*

$$J=\operatorname{diag}[\lambda_1 I_{s_1}+N_1, \cdots, \lambda_\mu I_{s_\mu}+N_\mu], \tag{1.1}$$

with nilpotent matrices $N_1, \cdots, N_\mu$ of the following type: For every ν, $1\leqq\nu\leqq\mu$, let l_ν denote the smallest positive integer for which

$$\operatorname{rank}(C-\lambda_\nu I)^{l_\nu}=s-s_\nu, \tag{1.2}$$

and define numbers $s_{\nu,k}$ by

$$s_{\nu,k}=\operatorname{rank}(C-\lambda_\nu I)^{l_\nu-k}-\operatorname{rank}(C-\lambda_\nu I)^{l_\nu-k+1}, \quad 1\leqq k\leqq l_\nu. \tag{1.3}$$

Then the numbers $s_{\nu,k}$ are (weakly) increasing in k, $1\leqq k\leqq l_\nu$, and if we block N_ν such that the kth diagonal block is of size $s_{\nu,k}$, $1\leqq k\leqq l_\nu$, we have $N_\nu=0$ if $l_\nu=1$, and otherwise

$$N_\nu=\begin{bmatrix} 0 & 0 & \cdot & 0 \\ M_{\nu,1} & 0 & \cdot & 0 \\ 0 & & \cdot & \cdot \\ \vdots & & \vdots & \vdots \\ 0 & \cdots & 0\, M_{\nu,l_\nu-1} & 0 \end{bmatrix}, \tag{1.4}$$

$$M_{\nu,k}=\underbrace{\begin{bmatrix} 0 & \cdot & \cdot & \cdot & 0 \\ \vdots & & & & \vdots \\ 0 & \cdot & \cdot & \cdot & 0 \\ 1 & 0 & \cdot & \cdot & 0 \\ 0 & \cdot & \cdot & \cdot & \cdot \\ \vdots & & & & \vdots \\ \cdot & \cdot & \cdot & \cdot & 0 \\ 0 & \cdot & \cdot & 0 & 1 \end{bmatrix}}_{s_{\nu,k}} \begin{array}{l} \left.\vphantom{\begin{matrix}0\\ \vdots\\ 0\end{matrix}}\right\} s_{\nu,k+1}-s_{\nu,k} \\ \left.\vphantom{\begin{matrix}1\\ 0\\ \vdots\\ 0\\ 0\end{matrix}}\right\} s_{\nu,k}\,, \end{array} \qquad 1\leqq k\leqq l_\nu-1\,. \tag{1.5}$$

Remark 1.1. In (1.5) it may occur that $s_{\nu,k+1}-s_{\nu,k}=0$ for some ν and k. In this case the reader should interprete (1.5) to read $M_{\nu,k}=I_{s_{\nu,k}}$. There will be other occasions during this paper, where an analogous interpretation has to be made in cases matrices are blocked with respect to a certain block structure where the number of rows and/or columns of certain blocks is zero.

Remark 1.2. Suppose natural numbers μ, l_ν, and $s_{\nu,k}$ $(1\leqq k\leqq l_\nu)$ with $s_{\nu,1}\leqq\cdots\leqq s_{\nu,l_\nu}$ $(1\leqq\nu\leqq\mu)$ are given such that $\sum_k s_{\nu,k}=s_\nu$, $\sum_\nu s_\nu=s$, and let J (with distinct numbers $\lambda_1,\cdots,\lambda_\mu$) be defined by (1.1), (1.4), (1.5). By calculating the powers of N_ν $(1\leqq\nu\leqq\mu)$ one can check that $N_\nu^k=0$ iff $k\geqq l_\nu$, and that

$$s_{\nu,k}=\operatorname{rank}(J-\lambda_\nu I)^{l_\nu-k}-\operatorname{rank}(J-\lambda_\nu I)^{l_\nu-k+1}\,, \qquad 1\leqq k\leqq l_\nu\,.$$

Hence if C is similar to J, then (1.2) and (1.3) follow. Therefore, given C it is sufficient to prove the existence of J which is similar to C, satisfying (1.1), (1.4), (1.5) for any natural μ, distinct $\lambda_1,\cdots,\lambda_\mu$, and natural numbers l_ν, $s_{\nu,1}\leqq\cdots\leqq s_{\nu,l_\nu}$ $(1\leqq\nu\leqq\mu)$ as above.

Remark 1.3. A short proof of Lemma 1 can be given by first putting C into Jordan canonical form and then applying a permutation similarity. However, we will give an induction type prove which does not use any other result than just matrix algebra and which also can be turned into an algorithm construction the transformation matrix taking C into its modified canonical form, once the eigenvalues of C are known.

Proof of Lemma 1. We proceed by induction with respect to s. For $s=1$, Lemma 1 holds trivially. Let now $s\geqq 2$. Without loss in generality, assume that C has zero as an eigenvalue (otherwise replace C by $C-\lambda I$ for some eigenvalue λ). By multiplication from the right with an invertible matrix T we can arrange that the first $\tilde{s}$ columns (with $\tilde{s}=\operatorname{rank} C<s$) of C are linearly independent whereas the

other columns are zero, and multiplication from the left with T^{-1} does not destroy this property. Hence we may directly assume

$$C=\begin{bmatrix} C_{11} & 0 \\ C_{21} & 0 \end{bmatrix}, \qquad C_{11} \text{ of size } \tilde{s}\times\tilde{s},$$

and by induction hypothesis (not that if $\tilde{s}=0$, the proof is completed) we may take

$$C_{11}=\tilde{J}=\operatorname{diag}[\lambda_1 \quad I_{\tilde{s}_1}+N_1, \cdots, \lambda_{\tilde{\mu}} \quad I_{\tilde{s}_{\tilde{\mu}}}+N_{\tilde{\mu}}],$$

with matrices N_ν satisfying (1.4), (1.5) corresponding to natural numbers $\tilde{l}_\nu$, $\tilde{s}_{\nu,1}\leqq \cdots \leqq \tilde{s}_{\nu,\tilde{l}_\nu}$ $(1\leqq\nu\leqq\tilde{\mu})$, $\sum_k \tilde{s}_{\nu,k}=\tilde{s}_\nu$, $\sum_\nu \tilde{s}_\nu=\tilde{s}$, and distinct numbers $\lambda_1, \cdots, \lambda_{\tilde{\mu}}$.

An application of a similarity transformation corresponding to a matrix of the form

$$\begin{bmatrix} I_{\tilde{s}} & 0 \\ T_{21} & I_{s-\tilde{s}} \end{bmatrix}$$

is equivalent to replacing C_{21} by $C_{21}-T_{21}\tilde{J}$. Hence if $\tilde{J}$ is invertible, we may take $T_{21}=C_{21}\tilde{J}^{-1}$, and the transformed C is then of the form

$$\begin{bmatrix} \tilde{J} & 0 \\ 0 & 0 \end{bmatrix}=J.$$

In view of Remark 1.2, this completes the proof of Lemma 1 if $\tilde{J}$ is invertible (with an obvious choice of the numbers μ, l_ν, $s_{\nu,k}$ and λ_ν).

If $\tilde{J}$ is not invertible we take $\mu=\tilde{\mu}$ and may arrange that $\lambda_\mu=0$. By computing $T_{21}\tilde{J}$ we see that we can arrange all the columns of $\tilde{C}_{21}=C_{21}-T_{21}\tilde{J}$ except for the last $\tilde{s}_{\mu,\tilde{l}_\mu}$ ones to be zero, i.e.

$$\tilde{C}_{21}=[0, \quad \tilde{C}]$$

with $\tilde{C}$ having $\tilde{s}_{\mu,\tilde{l}_\mu}$ columns that are linearly independent (otherwise C could not have had rank $\tilde{s}$; note that therefore $\tilde{s}_{\mu,\tilde{l}_\mu}\leqq s-\tilde{s}$, i.e. $\tilde{C}_{21}$ has at least as many rows as columns).

Finally, let $\tilde{T}=\operatorname{diag}[I_{\tilde{s}}, \hat{T}]$, $\hat{T}$ constant, invertible, of size $s-\tilde{s}$, and choose $\hat{T}$ such that its last $\tilde{s}_{\mu,\tilde{l}_\mu}$ columns equal $\tilde{C}$. Then

$$\tilde{T}^{-1}\begin{bmatrix} \tilde{J} & 0 \\ \tilde{C}_{21} & 0 \end{bmatrix}=\tilde{T}\begin{bmatrix} \tilde{J} & 0 \\ \hat{C}_{21} & 0 \end{bmatrix},$$
$$\hat{C}_{21}=[0, \hat{C}], \qquad \tilde{C}=\hat{T}\hat{C}.$$

By means of the above choice of $\hat{T}$, this implies

$$\hat{C}=\begin{bmatrix}0\\I\end{bmatrix}\}\tilde{s}_{\mu,\tilde{l}_\mu}=M_{\mu,\tilde{l}_\mu}\ .$$

Therefore, if we take

$$s_{\nu,k}=\tilde{s}_{\nu,k}\quad (1\leqq k\leqq l_\nu=\tilde{l}_\nu,\ 1\leqq\nu\leqq\mu-1),$$
$$l_\mu=\tilde{l}_\mu+1\ ,\qquad s_{\mu,k}=\tilde{s}_{\mu,k}\quad (1\leqq k\leqq l_\mu-1)\ ,\qquad s_{\mu,l_\mu}=s-\tilde{s}\ ,$$

then

$$J=\begin{bmatrix}\tilde{J} & 0\\ \hat{C}_{21} & 0\end{bmatrix}$$

satisfies (1.1), (1.4), (1.5), which completes the proof, using again Remark 1.2.

Remark 1.4. Up to an ordering of the eigenvalues, the modified canonical form is unique. This is another advantage over Jordan canonical form from the point of view taken in the theory of invariants; compare [1], Section 3a; [4], § 4a. It will be important later to actually think of an a-priori fixed ordering of the eigenvalues of J being made, so that we then can speak of *the unique modified canonical form* of a matrix.

Let $\mathcal{G}$ be a group (with respect to matrix multiplication) of constant, invertible matrices. We call $\mathcal{G}$ *an admissible group*, if $\mathcal{G}$ is the set of all the matrices C which, according to some fixed block structure $C=[C_{ij}]$, $1\leqq i,\ j\leqq m$, with square diagonal blocks, are lower triangularly blocked, invertible, and are characterized by (finitely many) conditions of the following two types:

1. For some pair (i, j), $1\leqq i,\ j\leqq m$, we have

(1.6) $$C_{ii}=C_{jj}\qquad \text{for every}\quad C\in\mathcal{G}\,.$$

2. For complex numbers α_{ij} $(1\leqq j<i\leqq m)$ we have

(1.7) $$\sum_{j<i}\alpha_{ij}C_{ij}=0\ ,$$

(with the implicit assumption that for those $\alpha_{ij}\neq 0$ the corresponding blocks C_{ij} are all of the same size, and that terms with $\alpha_{ij}=0$ may be omitted).

Remark 1.5. Note that there is only one block structure, according to which a group $\mathcal{G}$ can be admissible; we refer to it as *the associated block structure*. It is clear that not every set of invertible matrices defined by conditions of type 1 and 2 will be a group with respect to matrix multiplication (compare the following remark).

Remark 1.6. Let $\mathcal{G}$ be an admissible group, and think of $C=[C_{ij}]\in\mathcal{G}$ $(1\leqq i,$

$j \leqq m$) being blocked according to the associated block structure. Then we may divide $\{1, \cdots, m\}$ into disjoint subsets $\sigma_1, \cdots, \sigma_r$ such that the conditions of type one can altogether be expressed as

$$C_{ii}=C_{jj} \text{ if } i, j \text{ are in the same } \sigma_\nu, \qquad \text{for every } C \in \mathscr{G}.$$

Every condition (1.7) can then be rewritten as

$$\sum_{\mu=1}^{r} \sum_{\nu=1}^{r} \sum_{\substack{i,j \\ j<i, i \in \sigma_\nu, j \in \sigma_\mu}} \alpha_{ij} C_{ij} = 0 . \tag{1.8}$$

Since $\mathscr{G}$ contains *all* lower triangularly blocked invertible constant matrices satisfying the conditions of type 1, 2, we find $D \in \mathscr{G}$ whenever

$$D = \operatorname{diag}[D_1, \cdots, D_m], \quad D_i \text{ invertible } (i=1, \cdots, m),$$
$$D_i = D_j \qquad \text{if} \quad i, j \in \sigma_\nu \ (\nu=1, \cdots, r).$$

Therefore if $C \in \mathscr{G}$, then $DC \in \mathscr{G}$, i.e. (1.8) implies (if we denote the common value of $D_i = D_j$ for $i, j \in \sigma_\nu$ by D_ν, $\nu=1, \cdots, r$)

$$\sum_{\mu=1}^{r} \sum_{\nu=1}^{r} D_\nu \sum_{\substack{i,j \\ j<i, i \in \sigma_\nu, j \in \sigma_\mu}} \alpha_{ij} C_{ij} = 0 .$$

Since $D_1, \cdots, D_r$ can be arbitrary invertible matrices, we conclude

$$\sum_{\mu=1}^{r} \sum_{\substack{i,j \\ j<i, i \in \sigma_\nu, j \in \sigma_\mu}} \alpha_{ij} C_{ij} = 0 \qquad \text{for every } \nu, \quad 1 \leqq \nu \leqq r .$$

Applying an analogous argument to CD which is also in $\mathscr{G}$ shows that even

$$\sum_{\substack{i,j \\ j<i, i \in \sigma_\nu, j \in \sigma_\mu}} \alpha_{ij} C_{ij} = 0 \qquad \text{for every } (\nu, \mu); \quad 1 \leqq \nu, \quad \mu \leqq r .$$

Hence without loss in generality, we may only consider conditions (1.7) in which summation is restricted by $i \in \sigma_\nu$, $j \in \sigma_\mu$ for some pair (ν, μ); $1 \leqq \nu$, $\mu \leqq r$. This will be of importance later.

Lemma 2. *Let J be in modified canonical form, i.e. having all the properties listed in Lemma 1. Let $\mathscr{G}$ be the group of invertible matrices commuting with J. Then $\mathscr{G}$ is admissible.*

Proof. Let C be constant, invertible, such that

$$CJ = JC . \tag{1.9}$$

If (with the notations used in Lemma 1) we preliminarily block $C=[C_{jk}]$, $1 \leqq j$,

$k \leqq \mu$, such that the kth diagonal block is of size s_k, $1 \leqq k \leqq \mu$, it follows that C must be diagonally blocked. Hence it suffices to show that matrices commuting with a diagonal block of J form an admissible group, since a direct sum of admissible groups is again admissible. So without loss in generality, take $\mu=1$, $\lambda_1=0$,

$$J=N=\begin{bmatrix} 0 & \cdot & \cdot & 0 \\ M_1 & \cdot & \cdot & \cdot \\ 0 & \cdot & \cdot & \cdot \\ \vdots & & & \vdots \\ 0 \cdots 0 & & M_{l-1} & 0 \end{bmatrix}, \qquad M_k=\underbrace{\begin{bmatrix} 0 \\ I \end{bmatrix}}_{s_k}\begin{matrix}\}s_{k+1}-s_k \\ \}s_k\end{matrix}, \qquad 1 \leqq k \leqq l-1,$$

with natural numbers l, $s_1 \leqq s_2 \leqq \cdots \leqq s_l$, $\sum_k s_k = s$. If we now block $C=[C_{jk}]$ ($1 \leqq j, k \leqq l$) in the block structure of N, we see that (1.9) holds iff

$$C_{1,k+1}M_k=0\,, \qquad M_jC_{jl}=0\,, \qquad C_{j+1,k+1}M_k=M_jC_{jk} \quad (1 \leqq j,\ k \leqq l-1)\,. \tag{1.10}$$

Since $M_jX=0$ implies $X=0$, we conclude inductively for $k=l, l-1, \cdots, 2$ that $C_{jk}=0$ $(1 \leqq j<k)$; i.e. C is lower triangularly blocked in the block structure of J (which generally is *not* the block structure according to which $\mathcal{G}$ is admissible). Hence (1.10) reduces to

$$C_{j+1,k+1}M_k=M_jC_{jk}\,, \quad 1 \leqq k \leqq j \leqq l-1\,. \tag{1.11}$$

For fixed k, j $(1 \leqq k \leqq j \leqq l-1)$, if we block

$$C_{j+1,k+1}=\begin{bmatrix} C^{(1)} & C^{(2)} \\ \underbrace{C^{(3)}}_{s_{k+1}-s_k} & \underbrace{C^{(4)}}_{s_k} \end{bmatrix}\begin{matrix}\}s_{j+1}-s_j \\ \}s_j\end{matrix},$$

then (1.11) holds iff $C^{(2)}=0$, $C^{(4)}=C_{jk}$. If we now block C according to the finest possible block structure for which C still is lower triangularly blocked, then from the above discussion we see that $\mathcal{G}$ becomes admissible (this block structure is obtained by leaving the first diagonal block, and for $j \geqq 1$ dividing the $(j+1)$th diagonal block into one of size $s_{j+1}-s_j$ plus others having the same sizes and order as the ones into which the jth block was divided before).

Let now $G_m(z)$ be a formal meromorphic invariant of some meromorphic differential equation, i.e.

$$G_m(z)=z^JUe^{Q(z)}\,, \qquad J=J'+U'\,,$$

where

$$Q(z)=\operatorname{diag}[q_1(z)I_{s_1}, \cdots, q_l(z)I_{s_l}]$$

with positive integers $s_1, \cdots, s_l$ and distinct polynomials $q_1(z), \cdots, q_l(z)$ in a root of z which are ordered in a way such that $Q(z)$ shows a certain ***superblock structure*** (for details see [1], [4]), U and U' are explicitly given in terms of $Q(z)$, and J' is showing the same block structure as $Q(z)$, its blocks are equal within every superblock and taken to be in canonical form; here we take the modified canonical form defined in Lemma 1.

It was shown in [1], [4] that a constant invertible C commutes with G_m iff it commutes with $Q(z)$, R, and J', where R is built as a direct sum of superblocks, each of which is of the (blocked) form

$$\begin{bmatrix} 0 & 0 & \cdot & 0 & I \\ I & 0 & \cdot & 0 & 0 \\ 0 & \cdot & \cdot & \cdot & \cdot \\ \vdots & & & & \vdots \\ 0 & \cdot & 0 & I & 0 \end{bmatrix},$$

where all the blocks are of size s (which is also the common size of blocks in $Q(z)$ within this superblock). Therefore $\mathscr{G}_m$ consists of all constant invertible matrices that are diagonally blocked in the block structure of $Q(z)$ with equal blocks within every superblock and each block commuting with the corresponding block of J'. Using Lemma 2, we immediately obtain

Proposition 1. *The group $\mathscr{G}_m$ of constant invertible matrices that commute with a given meromorphic invariant $G_m(z)$ is an admissible group, if we take the modified canonical form instead of Jordan canonical form.*

It is this property of the group $\mathscr{G}_m$ that we will use in the next Sections.

2. $\mathscr{G}$-equivalence and admissible subgroups of $\mathscr{G}$

Let $\mathscr{G}$ be any admissible group of constant, invertible matrices of size n. Throughout this paper, let $P(z)$ always denote a matrix

$$P(z)=\begin{bmatrix} 1 & 0 & \cdot & 0 \\ p_{21}(z) & 1 & \cdot & 0 \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & 0 \\ p_{n1}(z) & \cdot & p_{n,n-1}(z) & 1 \end{bmatrix} \tag{2.1}$$

with polynomials $p_{ij}(z)$ in the variable z having zero constant term $(1 \leqq j < i \leqq n)$, and let K always denote a matrix

$$K = \operatorname{diag}[k_1, \cdots, k_n] \tag{2.2}$$

with integer entries k_j $(1 \leqq j \leqq n)$.

According to [1], Section 4a, or [4], page 52, we can uniquely decompose $P(z)z^K C$ for every $C \in \mathscr{G}$ as

$$P(z)z^K C = C(z)\tilde{P}(z)z^{\tilde{K}} \tag{2.3}$$

with $\tilde{P}(z)$, $\tilde{K}$ of the same type as $P(z)$, K, resp., and an analytic transformation $C(z)$, i.e. a matrix analytic at $z=\infty$ with $C(\infty)$ being invertible. If

$$P(z)z^K C = F(z) = [f_{ij}(z)], \tag{2.4}$$

then $\tilde{P}(z)z^{\tilde{K}}$ is calculated by a Gaussian algorithm type procedure, namely applying elementary row operations to $F(z)$ which correspond to multiplication from the left by analytic transformations.

Two matrices $P(z)z^K$ and $\tilde{P}(z)z^{\tilde{K}}$ as described above will be called *$\mathscr{G}$-equivalent,* if there is some $C \in \mathscr{G}$ such that (2.3) holds. Since $\mathscr{G}$ is a group, this indeed is an equivalence relation. The purpose of this paper is to describe an algorithmic procedure that selects a unique representative within every $\mathscr{G}$-equivalence class of matrices $P(z)z^K$ and gives a matrix $C \in \mathscr{G}$ that transforms a given $P(z)z^K$ into the corresponding representative by means of (2.3). This procedure will involve finitely many steps of the following type:

Given $P(z)$ and K, a matrix $C \in \mathscr{G}$ is constructed that transforms $P(z)z^K$ into $\tilde{P}(z)z^{\tilde{K}}$, such that whenever $P(z)z^K$ varies within a fixed $\mathscr{G}$-equivalence class, then $\tilde{P}(z)z^{\tilde{K}}$ will be seen to vary within a strictly smaller class of matrices which will turn out to be a $\hat{\mathscr{G}}$-equivalence class corresponding to an admissible subgroup $\tilde{\mathscr{G}}$ of $\mathscr{G}$. In order to ensure that the subgroups $\tilde{\mathscr{G}}$ which occur are actually admissible, we will make use of the following

Proposition 2. *Let $\mathscr{G}$ be any admissible group. If $C \in \mathscr{G}$ is blocked according to the associated block structure, then we define a subgroup of $\mathscr{G}$, say $\tilde{\mathscr{G}}$, by restricting the* kth *diagonal block of C (for some fixed k) to be taken from an admissible group $\hat{\mathscr{G}}$ of matrices of appropriate size. Then every such $\tilde{\mathscr{G}}$ is again admissible.*

Proof. With the same notation as in Remark 1.6, there is a unique ν_0 with $k \in \sigma_{\nu_0}$. By splitting the jth diagonal block (for every $j \in \sigma_{\nu_0}$) into subblocks according to the block structure associated with $\hat{\mathscr{G}}$, we define a new block structure according to which every $C \in \tilde{\mathscr{G}}$ is lower triangularly blocked. The new diagonal blocks are restricted by finitely many conditions of type 1 and otherwise free, whereas the lower triangular blocks which correspond to diagonal blocks in the original block structure are characterizable by finitely many conditions of type 2. Since every condition of the form (compare Remark 1.6)

$$\sum_{\substack{i,j \\ j<i,\, i\in\sigma_\nu,\, j\in\sigma_\mu}} \alpha_{ij} C_{ij}=0$$

(relative to the original block structure) can be replaced by finitely many conditions of the same type relative to the new block structure, this completes the proof.

Remark 2.1. Other ways of describing subgroups of an admissible group $\mathscr{G}$ that will play a role later are by either requiring that certain two diagonal blocks have to be equal or by defining an additional restriction (1.7). In the second case one has to ensure that the subset of matrices $C\in\mathscr{G}$ satisfying the additional equation actually forms a subgroup, but once this is done, then the subgroup clearly is admissible.

Before we start defining our normalization procedure, we wish to show that for a lower triangularly blocked matrix C, the problem of decomposing $P(z)z^K C$ according to (2.3) is mainly reduced to decomposing every one of the diagonal blocks. This can be seen from

Lemma 3. *Let* $C=\begin{bmatrix} C_{11} & 0 \\ C_{21} & C_{22}\end{bmatrix}$ *be a constant, invertible* $n\times n$ *matrix having square diagonal blocks, and let* $P(z)$ *and* K *be as in* (2.1), (2.2). *If we decompose* $P(z)z^K C$ *according to* (2.3), *then* $C(z)$ *is also lower triangularly blocked* (*like* C). *Furthermore, if we block* $\tilde{P}(z)$, $\tilde{K}$, $P(z)$, K *in the same block structure, then the diagonal blocks* $\tilde{P}_{jj}(z)z^{\tilde{K}_j}$ *of* $\tilde{P}(z)z^{\tilde{K}}$ *are found by decomposing* $P_{jj}(z)z^{K_j}C_{jj}$ *analogously to* (2.3) (*for* $j=1,2$), *and the off-diagonal block* $\tilde{P}_{21}(z)$ *is obtained by taking the polynomial part without constant terms of*

$$\tilde{P}_{22}(z)z^{\tilde{K}_2}C_{22}^{-1}z^{-K_2}P_{22}^{-1}(z)[P_{21}(z)z^{K_1}C_{11}+P_{22}(z)z^{K_2}C_{21}]z^{-\tilde{K}_1}\tilde{P}_{11}^{-1}(z)$$

and multiplying it from the right by $\tilde{P}_{11}(z)$.

Proof. The lower triangular block structure of $C(z)$ follows since every other matrix in (2.3) is blocked in the same way. Hence with obvious notations we receive by specializing (2.3) to the diagonal blocks

$$P_{jj}(z)z^{K_j}C_{jj}=C_{jj}(z)\tilde{P}_{jj}(z)z^{\tilde{K}_j} \qquad (j=1,2)\,. \tag{2.5}$$

This is a decomposition of $P_{jj}(z)z^{K_j}C_{jj}$ analogously to (2.3) and therefore determines $\tilde{P}_{jj}(z)z^{\tilde{K}_j}$ (since every such decomposition is unique). From the offdiagonal block in (2.3) we find

$$P_{21}(z)z^{K_1}C_{11}+P_{22}(z)z^{K_2}C_{21}=C_{21}(z)\tilde{P}_{11}(z)z^{\tilde{K}_1}+C_{22}(z)\tilde{P}_{21}(z)z^{\tilde{K}_1},$$

or equivalently

$$C_{22}^{-1}(z)C_{21}(z)+\tilde{P}_{21}(z)\tilde{P}_{11}^{-1}(z)=C_{22}^{-1}(z)[P_{21}(z)z^{K_1}C_{11}+P_{22}(z)z^{K_2}C_{21}]z^{-\tilde{K}_1}\tilde{P}_{11}^{-1}(z). \tag{2.6}$$

Since $\tilde{P}_{11}^{-1}(z)$ is of the same form as $\tilde{P}_{11}(z)$, we find that this means a decomposition of the right hand side into a sum of an analytic matrix (at $z=\infty$) plus a polynomial without constant term. This uniquely determines $\tilde{P}_{21}(z)\tilde{P}_{11}^{-1}(z)$, and solving (2.5) (for $j=2$) for $C_{22}^{-1}(z)$ and inserting it into (2.6) completes the proof.

Remark 2.2. Consider now any admissible group $\mathscr{G}$, and block $C\in\mathscr{G}$ with respect to the associated block structure. Then the last diagonal block, say C_{mm}, of C can be any constant invertible matrix. Furthermore, the last diagonal block of $\tilde{P}(z)z^{\tilde{K}}$ is determined by C_{mm} and the last diagonal block of $P(z)z^{K}$ alone. This motivates why in the next section we consider the group of all constant invertible matrices before then returning to the general case.

3. Normalisations in a special case

Throughout this section, let $\mathscr{G}$ be the group of all constant, invertible matrices of size n. Let $P(z)$, K be as in (2.1), (2.2). We first want to show that within any equivalence class with respect to $\mathscr{G}$-equivalence there are only a finite number of different matrices K occuring. Since for any fixed representative $P(z)z^{K}$, the matrix $\tilde{P}(z)z^{\tilde{K}}$ in (2.3) takes on every value within the equivalence class as C varies within $\mathscr{G}$, it suffices to prove the following estimates for the elements of $\tilde{K}$:

Lemma 4. *Let $P(z)z^{K}$ and $\tilde{P}(z)z^{\tilde{K}}$ of the above type be related by* (2.3) *for some $C\in\mathscr{G}$. Then*

$$\sum_{\nu} k_{\nu}-(n-1)M\leq\tilde{k}_j\leq M, \qquad 1\leq j\leq n, \tag{3.1}$$

where

$$M=\operatorname{Max}\{k_{\nu}+\deg p_{\mu\nu}\mid 1\leq\nu\leq\mu\leq n\} \qquad (p_{\mu\mu}(z)\equiv 1,\ 1\leq\mu\leq n). \tag{3.2}$$

Proof. In the usual way we define the degree (denoted by $\deg f$) of a function $f\not\equiv 0$ which is meromorphic at ∞ and extend the definition by $\deg f=-\infty$ if $f\equiv 0$. Then M is the maximal degree of the elements in $F(z)=P(z)z^{K}C$. Since multiplication with analytic transformations does not change the maximal degree (use that the inverse is also an analytic transformation), we find by solving (2.3) for $\tilde{P}(z)z^{\tilde{K}}$ that M is also the maximal degree of $\tilde{P}(z)z^{\tilde{K}}$. This proves the upper estimate in (3.1). To obtain the lower estimate, note that by taking determinants of both sides in (2.3) we find that $z^{\Sigma k_\nu}z^{-\Sigma \tilde{k}_\nu}$ is a (one-dimensional) analytic transformation, hence

$$\sum k_\nu = \sum \tilde{k}_\nu \leqq \tilde{k}_j + (n-1)M$$

for every $j=1, \cdots, n$.

Let now any $P(z)z^K$ be given. At the moment we are interested in two parameters associated with $P(z)z^K$; one is k_n (compare (2.2)), the other one is defined as the minimally chosen integer, say s, between 0 and $n-1$ such that

$$k_{s+1} = \cdots = k_n\,, \qquad \text{and} \qquad p_{ij}(z) \equiv 0 \quad (s+1 \leqq j < i \leqq n) \tag{3.3}$$

(with an obvious interpretation if $s=n-1$). As a first step towards a calculation of a representative we wish to characterize within every equivalence class those $P(z)z^K$ for which these two parameters are minimized in the sense that for every $\tilde{P}(z)z^{\tilde{K}}$ which is $\mathscr{G}$-equivalent to $P(z)z^K$ we have $\tilde{k}_n \geqq k_n$, and if $\tilde{k}_n = k_n$ then $\tilde{s} \geqq s$ (with analogous definition of $\tilde{s}$). Every such $P(z)z^K$ we call *minimal with respect to* (k_n, s), and Lemma 4 guarantees the existence of $P(z)z^K$ which are minimal in this sense.

In this context we may assume that $s \geqq 1$ (since for $s=0$ we have $K=kI_n$, $P(z) \equiv I_n$, and $\mathscr{G}$-equivalence reduces to equality in this case). For an arbitrarily given $P(z)z^K$ we distinguish two cases:

I). Suppose there exists a constant s-tupel $(c_1, \cdots, c_s) \neq (0, \cdots, 0)$ such that

$$\deg\Big(\sum_{j=1}^{i-1} p_{ij}(z)z^{k_j}c_j + z^{k_i}c_i\Big) \leqq k_n\,, \qquad 1 \leqq i \leqq s\,, \tag{3.4}$$

and

$$\deg\Big(\sum_{j=1}^{s} p_{ij}(z)z^{k_j}c_j\Big) \leqq k_n\,, \qquad s+1 \leqq i \leqq n\,. \tag{3.5}$$

Selecting $c_{s+1}, \cdots, c_n$ appropriately we then can arrange that

$$\deg\Big(\sum_{j=1}^{s} p_{ij}(z)z^{k_j}c_j + c_i z^{k_n}\Big) \leqq k_n - 1\,, \qquad s+1 \leqq i \leqq n\,.$$

If in (3.4) the "$<$" sign applies for $1 \leqq i \leqq s$, then take any $C=[c_{ij}] \in \mathscr{G}$ with $c_{in}=c_i$, $1 \leqq i \leqq n$. Using (3.3), we find for $F(z)=P(z)z^K C=[f_{ij}(z)]$, that

$$f_{in}(z) = \begin{cases} \displaystyle\sum_{j=1}^{i-1} p_{ij}(z)z^{k_j}c_j + z^{k_i}c_i\,, & 1 \leqq i \leqq s\,, \\ \displaystyle\sum_{j=1}^{s} p_{ij}(z)z^{k_j}c_j + c_i z^{k_n}\,, & s+1 \leqq i \leqq n\,. \end{cases} \tag{3.6}$$

Since $\tilde{k}_n$ always equals the maximal degree of $f_{in}(z)$ $(1 \leqq i \leqq n)$ (compare [4], page 52), we find in this case that $\tilde{k}_n \leqq k_n - 1$, hence $P(z)z^K$ is not minimal with respect to (k_n, s), and we replace the original $P(z)z^K$ by $\tilde{P}(z)z^{\tilde{K}}$ obtained from (2.3) for any choice of C as described above.

Now assume that equality holds in (3.4) at least for one index i_0, $1 \leqq i_0 \leqq s$. In this case we take any $C = [c_{ij}] \in \mathscr{G}$ with

$$\begin{aligned} c_{is} &= c_i && (1 \leqq i \leqq s)\,, \\ c_{is} &= 0 && (s+1 \leqq i \leqq n)\,, \\ c_{ij} &= \delta_{ij} && (1 \leqq i \leqq n,\ s+1 \leqq j \leqq n)\,. \end{aligned}$$

By following the decomposition procedure in [4], p. 52, we find for $\tilde{P}(z)z^{\tilde{K}}$ in (2.3):

$$\tilde{k}_s = \cdots = \tilde{k}_n = k_n\,, \qquad p_{ij}(z) \equiv 0 \quad (s \leqq j < i \leqq n)\,.$$

So in this case we replace $P(z)z^K$ by $\tilde{P}(z)z^{\tilde{K}}$ which has smaller value $\tilde{s} \leqq s-1$ and $\tilde{k}_n = k_n$.

We have seen so far that whenever we are in Case I, then $P(z)z^K$ is not minimal with respect to (k_n, s), and we have shown how to find a $C \in \mathscr{G}$ such that the modified $\tilde{P}(z)z^{\tilde{K}}$ is closer to being minimal in the sense that either $\tilde{k}_n < k_n$ or $\tilde{k}_n = k_n$, $\tilde{s} < s$.

II). Now suppose that the only s-tupel $(c_1, \cdots, c_s)$ for which (3.4) and (3.5) hold, is $c_1 = \cdots = c_s = 0$. We wish to show that then $P(z)z^K$ is minimal with respect to (k_n, s). Let $C \in \mathscr{G}$ be taken such that for $\tilde{P}(z)z^{\tilde{K}}$ obtained from (2.3) we find

$$\tilde{k}_n \leqq k_n\,, \qquad \text{and if} \quad \tilde{k}_n = k_n \quad \text{then} \quad \tilde{s} \leqq s\,.$$

If we define $c_i = c_{in}$, $1 \leqq i \leqq n$, then, as seen before, $\tilde{k}_n$ is the maximal degree of $f_{in}(z)$ $(1 \leqq i \leqq n)$ which are given by (3.6). So we see that $\tilde{k}_n \leqq k_n$ holds only if (3.4) and (3.5) are satisfied which according to our assumption implies $c_1 = \cdots = c_s = 0$, and consequently we have $\tilde{k}_n = k_n$. Therefore we see that k_n is minimal, and $\tilde{k}_n = k_n$ iff

$$c_{in} = 0\,, \qquad 1 \leqq i \leqq s\,.$$

Let now j_0, $s+1 \leqq j_0 \leqq n-1$, be taken such that

$$c_{ij} = 0\,, \qquad 1 \leqq i \leqq s\,, \quad j_0 + 1 \leqq j \leqq n$$

(if $s = n-1$, then this step may be omitted). We aim at showing that $\tilde{s} \leqq s$ implies $c_{ij_0} = 0$, $1 \leqq i \leqq s$. To do so, we take an invertible matrix $\hat{C}$ of size $n-s$ such that

$$C=\operatorname{diag}[I_s,\hat{C}]\tilde{C},$$

with $\tilde{C}=[\tilde{c}_{ij}]$ satisfying

$$\tilde{c}_{ij}=\delta_{ij}, \qquad 1\leqq i\leqq n, \quad j_0+1\leqq j\leqq n.$$

If we define $\hat{P}(z)z^{\hat{K}}$ by

$$P(z)z^K \operatorname{diag}[I_s,\hat{C}]=\hat{C}(z)\hat{P}(z)z^{\hat{K}}$$

with an analytic transformation $\hat{C}(z)$, then we easily see that $\hat{K}=K$, $P(z)=\operatorname{diag}[I_s,\hat{C}]\hat{P}(z)$, i.e. $\hat{C}(z)=\operatorname{diag}[I_s,\hat{C}]$. Consequently, since $\hat{P}(z)z^{\hat{K}}$ satisfies all the requirements on $P(z)z^K$ including that (3.4), (3.5) implies $c_1=\cdots=c_s=0$, we may without loss in generality assume that

$$c_{ij}=\delta_{ij}, \qquad 1\leqq i\leqq n, \quad j_0+1\leqq j\leqq n.$$

In this case we find

$$\tilde{k}_{j_0+1}=\cdots=\tilde{k}_n=k_n, \qquad \tilde{p}_{ij}(z)=0, \quad j_0+1\leqq j<i\leqq n.$$

We further see from the decomposition algorithm that $\tilde{k}_{j_0}$ equals the maximal degree of $f_{1j_0}(z),\cdots,f_{j_0j_0}(z)$, and that $\tilde{p}_{ij_0}(z)$ is the polynomial part (without constant term) of $f_{ij_0}(z)z^{-\tilde{k}_{j_0}}$ $(j_0+1\leqq i\leqq n)$. The $f_{ij_0}(z)$ are given by

$$f_{ij_0}(z)=\begin{cases}\sum_{j=1}^{i-1} p_{ij}(z)z^{k_j}c_{jj_0}+z^{k_i}c_{ij_0}, & 1\leqq i\leqq s,\\ \sum_{j=1}^{s} p_{ij}(z)z^{k_j}c_{jj_0}+z^{k_n}c_{ij_0}, & s+1\leqq i\leqq n.\end{cases}$$

Therefore $\tilde{k}_{j_0}=k_n$ and $\tilde{p}_{ij_0}(z)=0$ $(j_0+1\leqq i\leqq n)$ imply (3.4) and (3.5) with $c_j=c_{jj_0}$ $(1\leqq j\leqq s)$, which then implies $c_{jj_0}=0$ $(1\leqq j\leqq s)$.

The foregoing discussion shows that $\tilde{k}_n=k_n$ and $\tilde{s}\leqq s$ hold iff $c_{ij}=0$ $(1\leqq i\leqq s, s+1\leqq j\leqq n)$, and we finally show that then even $\tilde{s}=s$ holds. By the same argument as above, we may assume $c_{ij}=\delta_{ij}$ $(1\leqq i\leqq n,\ s+1\leqq j\leqq n)$, and in this case (with $j_0=s$) we conclude that $\tilde{s}<s$ is equivalent with (3.4) and (3.5) to hold for $c_j=c_{js}$ $(1\leqq j\leqq s)$, hence equivalent with $c_{js}=0$ $(1\leqq j\leqq s)$. But this contradicts to the invertibility of C, which proves $\tilde{s}=s$.

We formalize the results of the foregoing discussion as

Proposition 3. *Let $\mathscr{G}$ be the group of all constant invertible matrices of size $n(n\geqq 2)$, and let $P(z)z^K$ as in* (2.1), (2.2) *be given having parameter $s\geqq 1$. Then $P(z)z^K$ is minimal with respect to (k_n, s) iff the only s-tuple $(c_1,\cdots,c_s)$ for which*

$$\deg\Big(\sum_{j=1}^{i-1} p_{ij}(z)z^{k_j}c_j+z^{k_i}c_i\Big)\leqq k_n, \qquad 1\leqq i\leqq s,$$

and

$$\deg\left(\sum_{j=1}^{s} p_{ij}(z)z^{k_j}c_j\right) \leqq k_n\,, \qquad s+1 \leqq i \leqq n\,,$$

is $c_1 = \cdots = c_s = 0$.

In case of minimality, the matrices $C \in \mathscr{G}$ *which transform* $P(z)z^K$ *into* $\tilde{P}(z)z^{\tilde{K}}$ *with* $\tilde{k}_n = k_n$, $\tilde{s} = s$ *are precisely of the form*

$$C = \begin{bmatrix} C_{11} & 0 \\ C_{21} & C_{22} \end{bmatrix}, \tag{3.7}$$

where C_{11}, C_{22} *are arbitrary invertible matrices of size* s *and* $n-s$, *resp., and* C_{21} *is an arbitrary matrix of appropriate size.*

In case $P(z)z^K$ *is not minimal with respect to* (k_n, s) *then one can by a clear procedure described above construct a* $\mathscr{G}$*-equivalent matrix* $\tilde{P}(z)z^{\tilde{K}}$ *which is minimal.*

4. Normalizations in the general case

a) *Normalizations of the diagonal blocks*:

Let now $\mathscr{G}$ be any admissible group of $n \times n$ matrices, and consider any $P(z)z^K$ as in (2.1), (2.2). According to the block structure associated with $\mathscr{G}$ we block $P(z)$, K as

$$P(z) = \begin{bmatrix} P_{11}(z) & 0 & 0 \\ \vdots & & \vdots \\ \cdot & \cdot & 0 \\ P_{m1}(z) & \cdot & P_{mm}(z) \end{bmatrix}, \qquad P_{jj}(z) \text{ of size } s_j\,, \quad 1 \leqq j \leqq m\,, \tag{4.1}$$

$$K = \operatorname{diag}[K_1, \cdots, K_m]\,. \tag{4.2}$$

Let $C = [C_{ij}] \in \mathscr{G}$ be arbitrarily given and factor $P(z)z^K C$ according to (2.3). Since C is lower triangularly blocked, we see by means of Lemma 3 that the diagonal blocks $\tilde{P}_{jj}(z)z^{\tilde{K}_j}$ are obtained by decomposing $P_{jj}(z)z^{K_j}C_{jj}$ analogously to (2.3). For every fixed j, $1 \leqq j \leqq m$, the block C_{jj} can be any constant invertible matrix of appropriate size. Hence if we apply Proposition 3 to the jth diagonal block, we can replace $P(z)z^K$ by a $\mathscr{G}$-equivalent matrix (which we again denote by $P(z)z^K$) for which $P_{jj}(z)z^{K_j}$ is minimal (with respect to two parameters analogously defined as in Section 3). If $P_{jj}(z)z^{K_j} = z^{k_j}I_{s_j}$, then every $P(z)z^K$ has the minimality property. Hewever, if this is not the case, then those $P(z)z^K$ for which the jth diagonal block is minimal form a strictly smaller subset of the $\mathscr{G}$-equivalence class which we are working with, and from Propositions 2 and 3 we conclude that this subset is a $\tilde{\mathscr{G}}$-equivalence class with respect to an admissible subgroup $\tilde{\mathscr{G}}$ of $\mathscr{G}$.

A matrix $P(z)z^K$ is said to have *completely normalized diagonal blocks,* if

$$P_{jj}(z)z^{K_j}=z^{k_j}I_{s_j}\,,\qquad j=1,\cdots,m\,, \tag{4.3}$$

with integers k_j, $1\leqq j\leqq m$. If $P(z)z^K$ has completely normalized diagonal blocks, then every $\mathscr{G}$-equivalent matrix $\tilde{P}(z)$ has completely normalized diagonal blocks, and $\tilde{K}=K$. If some of the diagonal blocks of $P(z)z^K$ are not yet normalized, then by applying the foregoing discussion to anyone of the non-normalized diagonal blocks we construct a proper subset of the equivalence class of $P(z)z^K$ which is a $\tilde{\mathscr{G}}$-equivalence class for an admissible subgroup $\tilde{\mathscr{G}}$ of $\mathscr{G}$, and either every member of this subset has completely normalized diagonal blocks according to the finer block structure corresponding to $\tilde{\mathscr{G}}$ or we again choose a non-normalized diagonal block and construct another subset, and so forth. This *normalization procedure for the diagonal blocks* must within finitely many steps lead to a matrix having completely normalized diagonal blocks, since at every single *normalization step* the block structure associated with the subgroup $\tilde{\mathscr{G}}$ is strictly finer than the one associated with $\mathscr{G}$, and if the block sizes $s_j=1$, $1\leqq j\leqq m=n$, then (4.3) is trivially satisfied.

At every normalization step, we have chosen any diagonal block, which was not yet normalized, to work with. Since some diagonal blocks of $C\in\mathscr{G}$ may be restricted to be equal by means of conditions of type 1, the resulting finite sequence of equivalence classes and subgroups will in general depend upon the choice of the diagonal block. So in order to have a unique procedure, we think of an a-priori defined rule which decides which diagonal block we take in every normalization step. This rule must not depend upon the particular $P(z)z^K$ but only upon the $\mathscr{G}$-equivalence class which we are working with. As a natural example of such a *selection rule,* one may decide to take the diagonal block with maximal (minimal) index which is not normalized.

Theorem 1. *Let $\mathscr{G}$ be an admissible group, $P(z)z^K$ be any matrix as described in* (2.1), (2.2), *and blocked as in* (4.1), (4.2). *Then*

$$\tilde{K}=K$$

for every $\tilde{P}(z)z^{\tilde{K}}$ which is $\mathscr{G}$-equivalent to $P(z)z^K$ iff $P(z)z^K$ has completely normalized diagonal blocks. If $P(z)z^K$ does not have completely normalized diagonal blocks, then we may construct a $\mathscr{G}$-equivalent matrix $\tilde{P}(z)z^{\tilde{K}}$ and an admissible sub-

group $\tilde{\mathscr{G}}$ such that according to the block structure associated with $\tilde{\mathscr{G}}$ the diagonal blocks of $\tilde{P}(z)z^{\tilde{K}}$ are completely normalized. The construction yields unique $\tilde{P}(z)z^{\tilde{K}}$ and $\tilde{\mathscr{G}}$, if we proceed according to an arbitrary but fixed selection rule.

Proof. Most of the proof was given in the discussion preceding Theorem 1. It only remains to prove that $\tilde{K}=K$ (for every $\mathscr{G}$-equivalent $\tilde{P}(z)z^{\tilde{K}}$) implies that $P(z)z^K$ has completely normalized diagonal blocks. So we take any diagonal block of $P(z)z^K$ which for the sake of simplicity we may denote by $P(z)z^K$. If C denotes the corresponding diagonal block of an arbitrary matrix of $\mathscr{G}$, then C can be any constant invertible matrix. Suppose that $P(z)z^K \neq z^k I$, and that $P(z)z^K$ is minimal with respect to (k_n, s) (which implies $s \geqq 1$). Then we conclude from Proposition 3 that for $(c_1, \cdots, c_s)=(1, 0, \cdots, 0)$ either

$$\deg (z^{k_1}) > k_n\,,$$

or

$$\deg (p_{i1}(z)z^{k_1}) > k_n \qquad \text{for at least one } i\,, \quad 2 \leqq i \leqq n\,,$$

and if C is any constant, invertible matrix with $c_{1n}=1$, $c_{in}=0$ $(2 \leqq i \leqq n)$, then $\tilde{k}_n > k_n$, hence $\tilde{K} \neq K$.

Remark 4.1. Using Lemma 4, one finds another way to define a unique K among all the finitely many different K's occurring within a fixed equivalence class: Suppose that k_n is fixed to be the minimal possible value, then take k_{n-1} to be minimal with respect to the different possible values with fixed (minimal) k_n, and so forth. This might be considered a natural normalization, however the group of matrices relating any two $P(z)z^K$ and $\tilde{P}(z)z^{\tilde{K}}$ with "minimal" $K=\tilde{K}$ cannot so easily be determined and might turn out to be not admissible. In fact, one can make examples even when $n=3$, where this method of fixing K and the one explained in Theorem 1 are leading to different results.

b) *Normalization of the off-diagonal blocks.*

Suppose that (for any given admissible group $\mathscr{G}$) $P(z)z^K$ has completely normalized diagonal blocks, i.e. $P_{jj}(z) \equiv I_{s_j}$, $K_j = k_j I_{s_j}$, $1 \leqq j \leqq m$. Then for every $\tilde{P}(z)z^{\tilde{K}}$ which is $\mathscr{G}$-equivalent to $P(z)z^K$ we find $\tilde{K}=K$, $\tilde{P}_{jj}(z) \equiv I_{s_j}$, $1 \leqq j \leqq m$. The off-diagonal blocks of $\tilde{P}(z)$ and $P(z)$ can, however, still be different, and we will discuss now what we can do to normalize these blocks. For any (i, j), $1 \leqq j < i \leqq m$, the block $P_{ij}(z)$ is said to be *completely normalized,* if for every $\tilde{P}(z)z^K$ which is $\mathscr{G}$-equivalent to $P(z)z^K$

$$\tilde{P}_{ij}(z) = P_{ij}(z)\,. \tag{4.4}$$

Theorem 2. *Let $\mathscr{G}$ be an admissible group with an associated block structure having $m\geqq 2$ diagonal blocks, and let $P(z)z^K$ be blocked according to the same block structure and have all completely normalized blocks (including the diagonal blocks) except for $P_{m1}(z)$. Then for any $C\in\mathscr{G}$ we find when factoring $P(z)z^KC$ according to* (2.3)

$$C_{mm}\tilde{P}_{m1}(z)+C_{m1}(z)=P_{m1}(z)C_{11}+R_{m1}(z)\,, \tag{4.5}$$

where $R_{m1}(z)$ is meromorphic at ∞, and we can choose $C\in\mathscr{G}$ such that $\tilde{P}_{m1}(z)$ satisfies one of the following conditions:

a) *In case $\nu=\max\{\deg R_{m1}(z)\mid C\in\mathscr{G}\}\geqq 1$, we require $\tilde{P}_{m1}(z)$ not to contain a term z^ν.*

b) *In case $\deg R_{m1}(z)\leqq 0$ and $C_{mm}=C_{11}$ for every $C\in\mathscr{G}$, we require the highest coefficient of $\tilde{P}_{m1}(z)$ which is not a scalar multiple of the identity matrix to be in modified canonical form.*

c) *In case $R_{m1}(z)\leqq 0$ for every $C\in\mathscr{G}$, and C_{mm} can differ from C_{11} for certain $C\in\mathscr{G}$, we require the highest non-zero coefficient of $\tilde{P}_{m1}(z)$ either to be the identity matrix (if $P_{m1}(z)$ is square and has invertible highest coefficient) or otherwise to have the form*

$$\begin{bmatrix} 0 & 0\\ B & 0\end{bmatrix},$$

where B may be any square constant invertible matrix having the same size as the rank of the highest coefficient of $P_{m1}(z)$.

In everyone of these cases, the class of matrices $\tilde{P}(z)z^K$ which are $\mathscr{G}$-equivalent to $P(z)z^K$ and satisfy the normalizing condition form a $\tilde{\mathscr{G}}$-equivalence class with respect to an admissible subgroup $\tilde{\mathscr{G}}$ of $\mathscr{G}$.

Proof. As a direct consequence of the definition of admissible groups we see that if $C=[C_{ij}]\in\mathscr{G}$, then $\operatorname{diag}[C_{11},\cdots,C_{mm}]\in\mathscr{G}$. But for $\tilde{C}=\operatorname{diag}[C_{11},\cdots,C_{mm}]\in\mathscr{G}$ we obtain (using that $P(z)z^K$ has completely normalized diagonal blocks) that (2.3) holds iff $C(z)\equiv\tilde{C}$, hence $\tilde{C}\tilde{P}(z)=P(z)\tilde{C}$. Therefore (4.4) holds for every fixed $(i,j)\neq(m,1)$, $1\leqq j<i\leqq m$, iff either $P_{ij}(z)\equiv 0$ or $C_{ii}=C_{jj}$ for every $C\in\mathscr{G}$ and $P_{ij}=p_{ij}(z)I_s$ with $s=s_i=s_j$ and a scalar polynomial $p_{ij}(z)$ without constant term; hence in every case $P_{ij}C_{jj}=C_{ii}P_{ij}$ for $(i,j)\neq(m,1)$.

Comparing the $(m,1)$-block position of equation (2.3), we find

$$\sum_{i=1}^{m}P_{mi}(z)z^{k_iI_{s_i}}C_{i1}=\sum_{i=1}^{m}C_{mi}(z)\tilde{P}_{i1}(z)z^{k_1I_{s_1}}\,,$$

or equivalently (note that $C_{mm}(z)\equiv C_{mm}$, and use (4.4) for $i=2,\cdots,m-1$, $j=1$)

$$C_{mm}\tilde{P}_{m1}(z)+C_{m1}(z)=\sum_{i=1}^{m}P_{mi}(z)C_{i1}z^{k_i-k_1}-\sum_{i=2}^{m-1}C_{mi}(z)P_{i1}(z)\,. \tag{4.6}$$

Hence (4.5) holds with

$$R_{m1}(z)=\sum_{i=2}^{m}P_{mi}(z)C_{i1}z^{k_i-k_1}-\sum_{i=2}^{m-1}C_{mi}(z)P_{i1}(z)\,. \tag{4.7}$$

Using again (4.4) for $(i,j)\neq(m,1)$ and denoting the blocks of $P^{-1}(z)$ by $P_{ij}^{(-1)}(z)$, one obtains from (2.3)

$$C_{mi}(z)=\sum_{j=i}^{m}P_{mj}(z)\sum_{\nu=i}^{j}C_{j\nu}P_{\nu i}^{(-1)}(z)z^{k_j-k_\nu},\qquad 2\leqq i\leqq m-1\,.$$

Considering the terms $\nu=j$ separately, and using $P_{mj}C_{jj}=C_{mm}P_{mj}$ (compare above), we find

$$\begin{aligned}C_{mi}(z)&=C_{mm}\sum_{j=i}^{m}P_{mj}(z)P_{ji}^{(-1)}(z)+\sum_{j=i+1}^{m}P_{mj}(z)\sum_{\nu=i}^{j-1}C_{j\nu}P_{\nu i}^{(-1)}(z)z^{k_j-k_\nu}\\&=\sum_{j=i+1}^{m}P_{mj}(z)\sum_{\nu=i}^{j-1}C_{j\nu}P_{\nu i}^{(-1)}(z)z^{k_j-k_\nu}\,,\qquad 2\leqq i\leqq m-1\,.\end{aligned}$$

Hence it follows that

$$\begin{aligned}\sum_{i=2}^{m-1}C_{mi}(z)P_{i1}(z)&=\sum_{\nu=2}^{m-1}\sum_{j=\nu+1}^{m}P_{mj}(z)C_{j\nu}z^{k_j-k_\nu}\sum_{i=2}^{\nu}P_{\nu i}^{(-1)}(z)P_{i1}(z)\\&=-\sum_{\nu=2}^{m-1}\sum_{j=\nu+1}^{m}P_{mj}(z)C_{j\nu}P_{\nu 1}^{(-1)}(z)z^{k_j-k_\nu}\,.\end{aligned}$$

Inserting this into (4.7), we find that

$$R_{m1}(z)=\sum_{\nu=1}^{m-1}\sum_{j=\nu+1}^{m}P_{mj}(z)C_{j\nu}P_{\nu 1}^{(-1)}(z)z^{k_j-k_\nu}\,.$$

Since $\tilde{C}\tilde{P}(z)=P(z)\tilde{C}$ for every $\tilde{C}=\operatorname{diag}[C_{11},\cdots,C_{mm}]\in\mathscr{G}$ implies $P^{-1}(z)\tilde{C}=\tilde{C}\tilde{P}^{-1}(z)$ for every such $\tilde{C}$, we conclude (compare the beginning of the proof) that the non-zero blocks of $P^{-1}(z)$ (except for $P_{m1}^{(-1)}$) are scalar polynomials times identity matrices. Therefore we obtain

$$R_{m1}(z)=\sum_{\mu}z^{\mu}\sum_{\substack{i,j\\1\leq j<i\leq m}}\alpha_{ij}^{(\mu)}C_{ij}\,, \tag{4.8}$$

where the first sum is finite, and $\alpha_{ij}^{(\mu)}$ are complex numbers, which do not depend upon the particular $P(z)$ but only upon the $\mathscr{G}$-equivalence class we are working in (due to our assumptions).

Now suppose that the maximal degree of $R_{m1}(z)$ with respect to all $C\in\mathscr{G}$ is $\nu\geqq 1$. Then from (4.8) we learn that the highest coefficient of $R_{m1}(z)$ can take on every value, due to the structure of $\mathscr{G}$. Hence we may find $C\in\mathscr{G}$ such that $\tilde{P}_{m1}(z)$ does not contain a term z^{ν}. If the same holds for $P_{m1}(z)$ as well, then the subset $\tilde{\mathscr{G}}$ of all $C\in\mathscr{G}$ such that $\tilde{P}(z)$ satisfies normalization condition a) is precisely determined by

$$\sum_{\substack{i,j\\ 1\leq j<i\leq m}} \alpha_{ij}^{(\nu)} C_{ij}=0\,. \tag{4.9}$$

So if we can certify that $\tilde{\mathscr{G}}$ is a group, then it certainly is admissible. But since the coefficients $\alpha_{ij}^{(\nu)}$ do not depend upon the particular $P(z)$ (compare above), we see that $\tilde{\mathscr{G}}$ is independent of the particular $P(z)$ as long as it is normalized. Hence if $C_1, C_2\in\tilde{\mathscr{G}}$, then

$$P(z)z^K C_1=C_1(z)\tilde{P}(z)z^K\,,$$
$$\tilde{P}(z)z^K C_2=C_2(z)\hat{P}(z)z^K\,,$$

where $P(z)$, $\tilde{P}(z)$ and $\hat{P}(z)$ are normalized, and $C_1(z)$, $C_2(z)$ are analytic transformations, hence

$$P(z)z^K C_1C_2=C_1(z)C_2(z)\hat{P}(z)z^K\,,$$
$$\tilde{P}(z)z^K C_1^{-1}=C_1^{-1}(z)P(z)z^K\,,$$

which proves that both C_1^{-1} and C_1C_2 are in $\tilde{\mathscr{G}}$, so $\tilde{\mathscr{G}}$ is a group. This proves the Theorem in case a).

Suppose now $\deg R_{m1}(z)\leqq 0$ for every $C\in\mathscr{G}$. Then it follows from (4.5) that

$$\tilde{P}_{m1}(z)=C_{mm}^{-1}P_{m1}(z)C_{11}\,.$$

If $C_{mm}=C_{11}$ for every $C\in\mathscr{G}$, then $P_{m1}(z)$ is square and either $P_{m1}(z)=p_{m1}(z)I_{s_1}$ with a scalar polynomial $p_{m1}(z)$ (in which case even $P_{m1}(z)$ would be completely normalized), or we can normalize $\tilde{P}(z)$ by putting the highest non-scalar coefficient of $P_{m1}(z)$ into modified canonical form J. If we now assume that $P(z)$ had already been normalized, then $\tilde{P}(z)$ is normalized iff C satisfies

$$C_{11}J=JC_{11}\,. \tag{4.10}$$

Using Lemma 2 and Proposition 2, the group $\tilde{\mathscr{G}}$ of all $C\in\mathscr{G}$ which satisfy (4.10) is seen to be admissible, which completes the proof in case b).

Finally, if $\deg R_{m1}(z)\leqq 0$ for every $C\in\mathscr{G}$, and C_{mm} and C_{11} can vary independently, then we can arrange to normalize $\tilde{P}(z)$ by requiring that the highest non-zero coefficient of $P(z)$ either becomes the identity matrix (if it is square and invertible) or is brought into the form $\begin{bmatrix}0 & 0\\ B & 0\end{bmatrix}$, B square, invertible, but otherwise arbitrary. (Note that we do not have to consider $P_{m1}(z)\equiv 0$, since then $P_{m1}(z)$

would be completely normalized). Again, if $P(z)$ is assumed to be normalized, too, then the group $\tilde{\mathscr{G}}$ of all $C\in\mathscr{G}$ such that $\tilde{P}(z)$ is normalized is described by either $C_{11}=C_{mm}$ or by

$$C_{11}=\begin{bmatrix} C_{11}^{(1)} & 0 \\ C_{21}^{(1)} & C_{22}^{(1)} \end{bmatrix}, \qquad C_{mm}=\begin{bmatrix} C_{11}^{(m)} & 0 \\ C_{21}^{(m)} & C_{22}^{(m)} \end{bmatrix},$$

with square diagonal blocks and $C_{11}^{(1)}$, $C_{22}^{(m)}$ being of same size as B. In both cases $\tilde{\mathscr{G}}$ can be seen to be admissible, which completes the proof in case c).

In order to complete the definition of a procedure selecting a representative within every $\mathscr{G}$-equivalence class, we return to the point where an admissible group $\mathscr{G}$ together with any $P(z)z^K$ having completely normalized diagonal blocks are given. Select any pair of indices (i, j), $1\leqq j<i\leqq m$, such that for every $C\in\mathscr{G}$ that takes $P(z)z^K$ into $\tilde{P}(z)z^K$ we find

$$\tilde{P}_{\nu\mu}(z)=P_{\nu\mu}(z)\,, \qquad j\leqq\mu<\nu\leqq i\,, \quad (\nu,\mu)\neq(i, j)\,.$$

From (2.3) we then conclude

$$\begin{bmatrix} P_{jj}(z) & 0\cdots 0 \\ \vdots & \vdots \\ \cdot \quad \cdot & 0 \\ P_{ij}(z) \quad \cdot & P_{ii}(z) \end{bmatrix} \operatorname{diag}[z^{K_j},\cdots,z^{K_i}] \begin{bmatrix} C_{jj} & 0\cdots 0 \\ \vdots & \vdots \\ \cdot \quad \cdot & 0 \\ C_{ij} \quad \cdot & C_{ii} \end{bmatrix}$$
$$=\begin{bmatrix} C_{jj}(z) & 0\cdots 0 \\ \vdots & \vdots \\ \cdot \quad \cdot & 0 \\ C_{ij}(z) \quad \cdot & C_{ii}(z) \end{bmatrix} \begin{bmatrix} \tilde{P}_{jj}(z) & 0\cdots 0 \\ \vdots & \vdots \\ \cdot \quad \cdot & 0 \\ \tilde{P}_{ij}(z) \quad \cdot & \tilde{P}_{ii}(z) \end{bmatrix} \operatorname{diag}[z^{K_j},\cdots,z^{K_i}]\,.$$

In order to apply Theorem 2 (with i, j instead of $m, 1$), we have to ensure that the group of matrices

$$C^{(i,j)}=\begin{bmatrix} C_{jj} & 0\cdots 0 \\ \vdots & \vdots \\ \cdot \quad \cdot & 0 \\ C_{ij} \quad \cdot & C_{ii} \end{bmatrix}$$

is admissible. It is however clear that conditions requiring equality for certain diagonal blocks of C are either meaningless for $C^{(i,j)}$ or just restrict two of its diagonal blocks to be equal. So we consider the system of homogeneous equations restricting the lower triangular blocks. As we have seen in Remark 1.6, this system splits into a direct sum of smaller systems where the blocks involved are of fixed size, hence every block may be treated as one unknown. By considering a fixed one of those smaller systems and numbering the unknowns in a way that those corresponding to blocks in $C^{(i,j)}$ come last, we can by means of Gaussian algorithm find an equivalent system with upper triangular coefficient matrix. In this form, the equations coming last (if any) are those equations describing the structure of $C^{(i,j)}$ whereas the others are meaningless for $C^{(i,j)}$. This shows the admissibility of the group of $C^{(i,j)}$, hence all assumptions of Theorem 2 are satisfied (with obvious changes in notations) and we define a *normalizing step* to mean a step described in cases a), b), c) of Theorem 2, applied to the pair (i, j) instead of $(m, 1)$. By iterating this procedure, we finally come within finitely many steps to a point where $\mathscr{G}$-equivalence coincides with equality, since at every level the block structure associated to the admissible group becomes strictly finer and/or some of the parameters of the matrices $P(z)$ within the equivalence class are fixed.

Remark 4.2. Whenever we choose a pair (i, j) as described above, there may be several possible choices which lead to different normalizations. Hence analogously to the normalization of the diagonal blocks we have to define a *selection rule* in order to make the choice unique. Again, a natural example for such a rule may be to take j maximal and then with fixed j take i maximal such that the block $P_{ij}(z)$ is not yet completely normalized when $P(z)$ varies within a $\mathscr{G}$-equivalence class.

5. Applications

a) *An example*

Consider the case $n=3$, i.e. let

$$P(z)=\begin{bmatrix} 1 & 0 & 0 \\ p_{21}(z) & 1 & 0 \\ p_{31}(z) & p_{32}(z) & 1 \end{bmatrix}, \qquad K=\operatorname{diag}[k_1, k_2, k_3]$$

be arbitrarily given, and let $\mathscr{G}$ be the group of all invertible 3×3 matrices. Furthermore, suppose that $P(z)z^K$ is not a scalar multiple of I, since this would imply that $\mathscr{G}$-equivalence coincides with equality.

Applying the results of Sections 3 and 4a, one can determine a $\mathscr{G}$-equivalent matrix $\tilde{P}(z)z^{\tilde{K}}$ which has minimized parameters $\tilde{k}_3$ and $\tilde{s}$, and if $\tilde{s}=2$, one can also minimize $\tilde{k}_2$ (when $\tilde{k}_3$ and $\tilde{s}$ are kept minimal). Then if we assume that $P(z)z^K$ did already have those minimality properties, one easily determines the subgroup of $\mathscr{G}$ (which for the sake of simplicity may again be denoted by $\mathscr{G}$) such that the matrices $\tilde{P}(z)z^{\tilde{K}}$ which have these minimality properties are precisely the ones which are $\mathscr{G}$-equivalent to $P(z)z^K$. Doing this, one finds the following three different cases:

1. Suppose

$$\left.\begin{cases}\deg (z^{k_1}c_1)\leqq k_3 \\ \deg (p_{21}(z)z^{k_1}c_1+z^{k_2}c_2)\leqq k_3 \\ \deg (p_{31}(z)z^{k_1}c_1+p_{32}(z)z^{k_2}c_2)\leqq k_3\end{cases}\right\} \quad \text{implies} \quad c_1=c_2=0\,, \tag{5.1}$$

and

$$\left.\begin{cases}\deg (z^{k_1}c_1)\leqq k_2 \\ \deg (p_{21}(z)z^{k_1}c_1)\leqq k_2\end{cases}\right\} \quad \text{implies} \quad c_1=0\,. \tag{5.2}$$

Then

$$\mathscr{G}=\left\{\begin{bmatrix} c_{11} & 0 & 0 \\ c_{21} & c_{22} & 0 \\ c_{31} & c_{32} & c_{33}\end{bmatrix}, \quad c_{11}c_{22}c_{33}\neq 0\right\}.$$

2. Suppose $p_{21}(z)\equiv 0$, $k_1=k_2$, and (5.1) holds.

Then

$$\mathscr{G}=\left\{\begin{bmatrix} c_{11} & c_{12} & 0 \\ c_{21} & c_{22} & 0 \\ c_{31} & c_{32} & c_{33}\end{bmatrix}, \quad c_{33}\det\begin{bmatrix} c_{11} & c_{12} \\ c_{21} & c_{22}\end{bmatrix}\neq 0\right\}.$$

3. Suppose $p_{32}(z)\equiv 0$, $k_2=k_3$, and

$$\left.\begin{cases}\deg (z^{k_1}c_1)\leqq k_3 \\ \deg (p_{21}(z)z^{k_1}c_1)\leqq k_3 \\ \deg (p_{31}(z)z^{k_1}c_1)\leqq k_3\end{cases}\right\} \quad \text{implies} \quad c_1=0\,. \tag{5.3}$$

Then

$$\mathscr{G}=\left\{\begin{bmatrix} c_{11} & 0 & 0 \\ c_{21} & c_{22} & c_{23} \\ c_{31} & c_{32} & c_{33}\end{bmatrix}, \quad c_{11}\det\begin{bmatrix} c_{22} & c_{23} \\ c_{32} & c_{33}\end{bmatrix}\neq 0\right\}.$$

It is now easy to apply the results of Section 4b to see what further normalizations can be made. We indicate this procedure in Case 1: Since every $C\in\mathscr{G}$ is now lower triangular, we find that (2.3) implies $\tilde{K}=K$, $C(z)$ lower triangular, and $c_{ii}(z)=c_{ii}$ $(i=1, 2, 3)$. Using this, we find three equations determining $\tilde{P}(z)z^K$ for every $C\in\mathscr{G}$.

(5.4) $$c_{33}\tilde{p}_{32}(z)+c_{32}(z)=p_{32}(z)c_{22}+z^{k_3-k_2}c_{32}\,,$$

(5.5) $$c_{22}\tilde{p}_{21}(z)+c_{21}(z)=p_{21}(z)c_{11}+z^{k_2-k_1}c_{21}\,,$$

(5.6) $$c_{33}\tilde{p}_{31}(z)+c_{31}(z)=p_{31}(z)c_{11}+p_{32}(z)z^{k_2-k_1}c_{21}+z^{k_3-k_1}c_{31}-c_{32}(z)\tilde{p}_{21}(z)\,.$$

(Compare this to (4.5)). Whenever $k_3-k_2\geqq 1$, we may arrange that $\tilde{p}_{32}(z)$ does not contain a term $z^{k_3-k_2}$, and the set of all matrices $P(z)z^K$ (within our equivalence class) for which $p_{32}(z)$ does not contain a term $z^{k_3-k_2}$ is precisely an equivalence class with respect to the subgroup of matrices C with $c_{32}=0$. If $k_3-k_2\leqq 0$, then (5.4) implies

(5, 4′) $$\begin{cases} c_{32}(z)=z^{k_3-k_2}c_{32}\,, \\ c_{33}\tilde{p}_{32}(z)=p_{32}(z)c_{22}\,, \end{cases}$$

and the same equations hold in the case $k_3-k_2\geqq 1$ if we restrict to those C with $c_{32}=0$. So we see next, that either $p_{32}(z)\equiv 0$ (in which case $\tilde{p}_{32}(z)\equiv 0$) or we may normalize $\tilde{p}_{32}(z)$ to have highest coefficient one. Then the subset of $P(z)z^K$ where $p_{32}(z)$ has highest coefficient one is an equivalence class corresponding to the subgroup of C with $c_{22}=c_{33}$ (which again is admissible).

Quite the same arguments show how to normalize $p_{21}(z)$, and once all the possible normalizations of $p_{32}(z)$ and $p_{21}(z)$ are carried out, we are left with equation (5.6). But since now $\tilde{p}_{21}(z)=p_{21}(z)$ and $c_{32}(z)=z^{k_3-k_2}c_{32}$, it is quite easy to see how we can normalize $p_{31}(z)$.

b) *Canonical forms of constant matrices with respect to a restricted type of similarity*:

Suppose that a group $\mathscr{G}$ of constant, invertible $n\times n$ matrices is given. We call two arbitrary constant $n\times n$ matrices A, B $\mathscr{G}$-similar, if for some $C\in\mathscr{G}$

(5.7) $$AC=CB\,,$$

and one might ask for a canonical form of a matrix A under $\mathscr{G}$-similarity. In case $\mathscr{G}$ is admissible, we can handle this problem: Given any $n\times n$ matrix A, form a $2n\times 2n$ matrix $P(z)$ by

(5.8) $$P(z)=\begin{bmatrix} I_n & 0 \\ zA & I_n \end{bmatrix},$$

and for $C\in\mathscr{G}$, let

(5.9) $$\tilde{C}=\operatorname{diag}[C, C]\,.$$

Then one finds that the set $\tilde{\mathscr{G}}$ of $\tilde{C}$ with $C\in\mathscr{G}$ is an admissible group, and if (with $K=0$) we decompose $P(z)\tilde{C}$ into $C(z)\tilde{P}(z)z^{\tilde{K}}$ according to (2.3), we find

$$\tilde{K}=0\,,\qquad C(z)=\tilde{C}\,,$$
$$\tilde{P}(z)=\begin{bmatrix} I_n & 0 \\ zB & I_n \end{bmatrix},\qquad B=C^{-1}AC\,.$$

Hence we see that the problem of finding a representative for $P(z)$ with respect to $\tilde{\mathscr{G}}$-equivalence is completely the same as finding a canonical form for A with respect to $\mathscr{G}$-similarity. The procedure defined in this paper therefore can be used to calculate such a canonical form for every given A and every admissible group $\mathscr{G}$, and in case $\mathscr{G}$ consists of arbitrary invertible matrices, then this canonical form coincides with the modified canonical form defined in Section 1.

References

[1, 2, 3] W. Balser, W. B. Jurkat, and D. A. Lutz: *A general theory of invariants for meromorphic differential equations*; *Part I, formal invariants*, Funk. Ekvac. **22** (1979) 197-221; *Part II, proper invariants*, Funk. Ekvac. **22** (1979) 257-283; *Part III*, Houston J. Math. **6** (1980) 149-189.

[4] W. B. Jurkat: *Meromorphe Differentialgleichungen*, Lecture Notes in Math. Vol. 637 (entire volume), Springer Verlag, Berlin-Heidelberg-New York, 1977.

Universität Ulm
Abt. f. Mathematik
Oberer Eselsberg
7900 Ulm
West Germany