

REDUCTION OF ORDINARY DIFFERENTIAL EQUATIONS TO THE BIRKHOFF CANONICAL FORM

BY
H. L. TURRITTIN⁽¹⁾

1. **Introduction.** In 1909 and again in 1913 G. D. Birkhoff attempted to reduce all homogeneous linear ordinary differential equations with single-valued analytic coefficients to a certain canonical form, (see [1 and 2] at the end of this paper). In 1953 F. R. Gantmacher [3] and again in 1959 P. Masani [4] pointed out by means of counter-examples that Birkhoff's 1913 theorem (reference [2, p. 136]) regarding the reduction to canonical form was false. Indeed both of Birkhoff's papers are very sketchy and of dubious validity when the monodromic matrix has multiple elementary divisors.

Our objective in this paper is to point out exactly where Birkhoff goes wrong and to decide what his procedure really does prove when the monodromic matrix has multiple elementary divisors. The author had hoped to push the analysis still further and determine the degree of simplicity that one might obtain in the canonical form when nonsingular transformations of various types are used in making the reduction. He has however been able so far to make only two steps in this direction, see Theorems III and IV. Much work remains to be done on this problem.

2. **A revised canonical form.** To be more precise consider a linear matrix differential equation

$$(1) \quad \frac{dX}{d\tau} = \tau^q A(\tau)X,$$

where X and A are square matrices with n rows and n columns; q is any integer and the descending power series

$$(2) \quad A = A(\tau) = \sum_{k=0}^{\infty} A_k/\tau^k \quad (A_0 \neq 0)$$

converges for $\tau_0 \leq |\tau| \leq \infty$. One would expect that equation (1) could be simplified and standardized by means of a transformation of the type

$$(3) \quad X = B(\tau)Y,$$

Received by the editors February 21, 1962 and, in revised form, June 25, 1962.

(1) The author has prepared this paper while in Austria on sabbatical leave from the University of Minnesota and supported half-time by a grant No. G 14879 from the National Science Foundation.

where the n^2 elements in matrix B are all analytic in some neighborhood of infinity, $\tau_0 \leq \tau_1 \leq |\tau| \leq \infty$, and the determinant of $B(\infty)$ is not zero.

THEOREM I. *There exists an integer s and a transformation of type (3) with $B(\infty) = I$, where I is the identity matrix, which will convert equation (1) into a canonical equation of the form*

$$(4) \quad \frac{dY}{d\tau} = (\tau^q) \sum_{j=0}^s C_j Y / \tau^j,$$

where C_0, \dots, C_s are appropriate constant square matrices.

The size of the integer s is a measure of the simplicity of this canonical form, the smaller the s the simpler the equation. Birkhoff claimed [2] that in no case would it be necessary to take s greater than $(q+1)$ if B were chosen properly. However in Gantmacher's counter-example, where $q = -1$, for instance, it turns out that it is necessary to take $s = 1$.

The proof of Theorem I, which is about to be given, is based on

THEOREM II. *Let $L(\tau)$ be a square matrix of functions single-valued and analytic for $\tau_1 \leq |\tau| < \infty$ and such that the determinant of L does not vanish for $\tau_1 \leq |\tau| < \infty$. Then there exists a matrix $B(\tau)$ of functions analytic at $\tau = \infty$, reducing to the unit matrix I at ∞ , and a matrix $N(\tau)$ of entire functions of determinant nowhere zero in the finite complex τ -plane, such that*

$$L(\tau) = B(\tau)N(\tau)K(\tau),$$

where K is a suitable diagonal matrix

$$(5) \quad K(\tau) = D\{\tau^{\alpha_1}, \tau^{\alpha_2}, \dots, \tau^{\alpha_n}\}$$

and $\alpha_1, \dots, \alpha_n$ are integers.

The symbolism in (5), where D is followed by braces, is used throughout this paper to indicate a diagonal matrix and the matrix elements on the main diagonal are exhibited within the braces and listed in order as they appear running down the main diagonal.

Birkhoff gives a valid proof of Theorem II in reference [5] and later points out, (see [6]), that his proof of Theorem II is a special case of an important theorem due to D. Hilbert [7] and J. Plemelj [8].

From the basic existence theorem for equations of type (1) we infer there exists a fundamental matrix solution $X = X_f(\tau)$ of (1) with elements which are analytic in some neighborhood of infinity; i.e., in $\tau_0 \leq \tau_2 \leq |\tau| < \infty$. The elements however are usually not single-valued functions of τ .

When $X_f(\tau)$ is continued analytically once clockwise around a simple closed circuit enclosing $\tau = 0$ and keeping $|\tau| > \tau_2$, the solution X_f becomes equal to a new fundamental solution

$$(6) \quad \tilde{X}_f(\tau) = X_f(\tau)J,$$

where the X_f can be initially so chosen that the constant monodromic matrix J will have the Jordan classical canonical form

$$(7) \quad J = D\{\rho_1 I_1, \dots, \rho_\alpha I_\alpha; \rho_{\alpha+1} I_{\alpha+1} + E_{\alpha+1}, \dots, \rho_\beta I_\beta + E_\beta\},$$

where the ρ 's are the characteristic roots of J and are nonvanishing, since \tilde{X}_f is a fundamental solution. The I_1, \dots, I_β are identity matrices of order n_1, \dots, n_β with $n_1 + \dots + n_\beta = n$ and the $E_{\alpha+1}, \dots, E_\beta$ are square matrices of respective orders $n_{\alpha+1}, \dots, n_\beta$. Every element in a matrix $E_{\alpha+1}, \dots, E_\beta$ is zero, save for the 1's running down the first subdiagonal of each of these E -matrices. The reader will note the notation introduced in (5) is being generalized to cover diagonal matrices with elements which are themselves submatrices.

To take care of complicated monodromic matrices, it is convenient to introduce the symbol $T\{a_1, a_2, \dots, a_n\}$ to denote a "triangular" or T -matrix, i.e., an n by n square matrix with a_1 appearing in every position on the main diagonal, zeros above the main diagonal, a_2 appearing in every position on the first subdiagonal, a_3 in every position on the second subdiagonal, etc.

Consider a T -matrix

$$F(\tau, \rho, n) = \tau^\lambda T\{1, \gamma_2(\tau, \rho, 1), \dots, \gamma_n(\tau, \rho, 1)\}$$

of n rows and let $\gamma_1(\tau, \rho, x) \equiv 1$. Then, writing ρ in its polar form $\rho = re^{i\theta}$, $0 \leq \theta < 2\pi$, $r = |\rho|$ we take

$$(8) \quad \lambda = -(\log \rho)/2\pi i = -(\theta - i \log r)/2\pi$$

and set

$$(9) \quad \gamma_j(\tau, \rho, x) = \frac{(-1)^{j-1}}{(j-1)!} \prod_{k=1}^{j-1} \frac{\log \tau}{2\pi i \rho} + \left(\frac{k-x}{\rho} \right) \quad \text{if } j = 2, \dots, n.$$

Note especially, because of the presence of $\log \tau$, that, when the independent variable τ varies continuously and makes a single clockwise circuit around the origin $\tau = 0$, the matrix $F(\tau, \rho, n)$ becomes equal to a new matrix

$$(10) \quad \tilde{F}(\tau, \rho, n) = F(\tau, \rho, n)(\rho I + E),$$

where I is the identity matrix of order n and $E = T\{0, 1, 0, \dots, 0\}$ is a T -matrix of order n .

To see that (10) is correct one needs the identity

$$(11) \quad \gamma_j(\tau, \rho, 2) - \gamma_j(\tau, \rho, 1) = \gamma_{j-1}(\tau, \rho, 1)/\rho \quad (j = 2, \dots, n)$$

and (11) follows at once from the fact that the left member of this identity can be written in the form

$$\frac{(-1)^{j-1}}{(j-1)!} \prod_{k=2}^{j-1} b_{k-2} \{b_{-1} - b_{j-2}\} = \frac{(-1)^{j-2}}{\rho(j-2)!} \prod_{k=1}^{j-2} b_{k-1},$$

using the abbreviation

$$(12) \quad b_h = \{(\log \tau)/2\pi i + h\}/\rho,$$

and noting that $b_{-1} - b_{j-2} = (1-j)/\rho$.

If one defines matrix $L(\tau)$ by means of the formula

$$(13) \quad X_f(\tau) = L(\tau)M(\tau),$$

where

$$M(\tau) = D\{\tau^{\lambda_1} I_1, \dots, \tau^{\lambda_\alpha} I_\alpha; F(\tau, \rho_{\alpha+1}, n_{\alpha+1}), \dots, F(\tau, \rho_\beta, n_\beta)\};$$

and $\lambda_j = (-\log \rho_j)/2\pi i$, ($j = 1, \dots, \beta$), selecting for the logarithm the value indicated in (8), then the matrix L will have all the properties of the matrix L in Theorem II. This statement is true because M has been so constructed that when the variable τ makes a clockwise circuit once around the origin, the continuously varying matrix $M(\tau)$ becomes equal to $\tilde{M}(\tau) = M(\tau)J$. Hence using the notation of Theorem II, the fundamental solution X_f takes the form

$$(14) \quad X_f(\tau) = B(\tau)N(\tau)K(\tau)M(\tau).$$

As suggested by Birkhoff, apply a transformation of type (3) to equation (1), selecting for the $B(\tau)$ in (3) the matrix $B(\tau)$ in (14) and thus obtain a new equation

$$(15) \quad \frac{dY}{d\tau} = C(\tau)Y,$$

where

$$(16) \quad C = \tau^q B^{-1}AB - B^{-1} \frac{dB}{d\tau}.$$

Observe that, since B has a convergent representation

$$B(\tau) = I + \sum_{j=1}^{\infty} B_{ji}/\tau^j$$

when $|\tau|$ is sufficiently large, it follows from (16) that C also has a convergent representation

$$(17) \quad C(\tau) = \tau^q \sum_{j=0}^{\infty} C_j / \tau^j,$$

when $|\tau|$ is sufficiently large.

Moreover one notes that equation (15) has a fundamental solution

$$Y_f = N(\tau)K(\tau)M(\tau).$$

Hence also

$$(18) \quad C = \frac{dY_f}{d\tau} Y_f^{-1} = \frac{dN}{d\tau} N^{-1} + N \frac{dK}{d\tau} K^{-1} N^{-1} + NK \frac{dM}{d\tau} M^{-1} K^{-1} N^{-1}.$$

As regards the three terms on the right in (18), one observes first that both $dN/d\tau$ and N^{-1} are matrices of entire functions of τ and thus the product $(dN/d\tau) N^{-1}$ is also a matrix of entire functions of τ .

In view of (5)

$$\frac{dK}{d\tau} K^{-1} = \tau^{-1} D\{\alpha_1, \alpha_2, \dots, \alpha_n\}.$$

Therefore $N(dK/d\tau)K^{-1}N^{-1}$ is a matrix of entire functions divided by τ .

If the E 's in (7) were missing, it would follow that

$$M = D\{\tau^{\lambda_1} I_1, \dots, \tau^{\lambda_\beta} I_\beta\} \text{ and } K \frac{dM}{d\tau} M^{-1} K^{-1} = D\{\lambda_1 I_1, \dots, \lambda_\beta I_\beta\} / \tau$$

and thus the last term of (18) would also be a matrix of entire functions divided by τ . In this event $C(\tau)$ would be a matrix of entire functions divided by τ and, as a consequence of this last fact and (17), it would be evident that the C in (15) would have the structure exhibited in (4) with $s = q + 1$, as predicted by Birkhoff.

However, if the E 's are present the evaluation of $(dM/d\tau)M^{-1}$ is more involved and the following result is needed:

$$(19) \quad \frac{dF(\tau, \rho, n)}{d\tau} = \{G(\rho, n) + \lambda I\} F(\tau, \rho, n) / \tau,$$

where the matrix

$$G(\rho, n) = T \left\{ 0, -1, \frac{1}{2\rho}, -\frac{1}{3\rho^2}, \dots, \frac{(-1)^{n-1}}{(n-1)\rho^{n-2}} \right\} / 2\pi i \rho.$$

In order to verify (19) we remark first that the product of two T -matrices is always a T -matrix. Equation (19) will follow at once if one can show that for $j = 2, \dots, n$

$$(20) \quad \frac{d\gamma_j(\tau, \rho, 1)}{d\tau} = \sum_{v=2}^j (-1)^{v-1} \gamma_{j-v+1}(\tau, \rho, 1) / 2\pi i \rho (v-1) \rho^{v-2}.$$

It is easy to verify (20) if $j = 2$. To proceed by induction, assume (20) is valid as it stands and we wish to show that (20) is correct if in (20) the j is replaced by $j + 1$. Since in an abbreviated notation, see (12),

$$(21) \quad \gamma_{j+1} = -b_{j-1}\gamma_j/j \quad (j = 1, 2, \dots, n-1)$$

it follows from the hypothesis for the induction that

$$\begin{aligned} \frac{d\gamma_{j+1}}{d\tau} &= \left(-b_{j-1} \frac{d\gamma_j}{d\tau} - \gamma_j/2\pi i \rho \tau \right) / j \\ &= - \left[\gamma_j + \sum_{v=2}^j (-1)^{v-1} \gamma_{j-v+1} \{ b_{j-v} + (v-1)/\rho \} / (v-1) \rho^{v-2} \right] / 2\pi i \rho \tau j. \end{aligned}$$

Hence by virtue of (21)

$$\frac{d\gamma_{j+1}}{d\tau} = - \left[\gamma_j + \sum_{v=2}^j (-1)^{v-1} \{ (v-j-1) \gamma_{j-v+2} + (v-1) \gamma_{j-v+1} / \rho \} \right] / 2\pi i \rho \tau j$$

and, combining like terms, one obtains (20) with the j replaced by $j + 1$. Therefore the induction is complete and we know (20) is valid for $j = 2, \dots, n$.

Thus (19) is substantiated and consequently

$$\frac{dF}{d\tau} F^{-1} = \{ \lambda I + G(\rho, n) \} / \tau.$$

Applying this formula to each of the diagonal blocks, the much needed result

$$\begin{aligned} \frac{dM}{d\tau} M^{-1} &= \tau^{-1} D \{ \lambda_1 I_1, \dots, \lambda_\alpha I_\alpha; \lambda_{\alpha+1} I_{\alpha+1}, \dots, \lambda_\beta I_\beta \} \\ &\quad + \tau^{-1} D \{ 0, \dots, 0; G(\rho_{\alpha+1}, n_{\alpha+1}), \dots, G(\rho_\beta, n_\beta) \} \end{aligned}$$

is procured.

Returning to (18) to select the appropriate s in (4) and remembering that (17) still holds, it is clear the G 's make complications and in general destroy the correctness of Birkhoff's theorem [2, p. 136]. To be more precise rewrite K in the form

$$K = D \{ K_1, \dots, K_\alpha; K_{\alpha+1}, \dots, K_\beta \},$$

where K_i has a structure (5), assuming in (5) we have replaced $\alpha_1, \dots, \alpha_n$ by an appropriate set of integers $\alpha_{i1}, \dots, \alpha_{i, n_i}$. The n_i denotes the number of rows in the matrix I_i in (7), ($i = 1, \dots, \beta$). It is now evident *the s in (4) is determined by the lowest negative integral power of τ , say τ^{-s_0} , which appears in the nonzero elements in the various matrix products*

$$K_i G(\rho_i, n_i) K_i^{-1} \quad (i = \alpha + 1, \dots, \beta).$$

In other words the s in (4) equals $q + 1 + s_0$.

If by chance

$$(22) \quad \alpha_{i1} \leq \alpha_{i2} \leq \cdots \leq \alpha_{i, n_i} \quad \text{for } i = \alpha + 1, \dots, \beta,$$

one sees $s_0 = 0$ and $s = q + 1$, again making Birkhoff's analysis right, but not otherwise. Since Theorem II gives no special information regarding the value of the integers $\alpha_1, \dots, \alpha_n$ in (5), it would appear that s_0 might turn out to be almost any integer. This discussion completes the proof of Theorem I and, as far as the author can see, no sharpening of the result is possible if the analysis is based only on Theorem II.

3. A more general type of reduction. It would appear that the minimal value of s in (4), which is to be associated with a particular given equation of type (1), can be determined in general only by an analysis that takes into account the more detailed structure of the A_k matrices in (2). The author has a definite result in this direction only in the case $q = -1$. This result will be presented in §4.

For the time being, however, let us return to the point of view taken by Birkhoff in his 1909 paper [1]. There an attempt was made to reduce an equation of type (1) to form (4) with $s = q + 1$ by means of a more general type of transformation

$$(23) \quad X = V(\tau)Y,$$

where V can be represented by a convergent series of the form

$$(24) \quad V(\tau) = \tau^p \sum_{j=0}^{\infty} V_j / \tau^j, \quad V_0 \neq 0, \quad 0 < r_0 < |\tau| < \infty;$$

where p is a suitably chosen integer and the determinant of $V(\tau)$ does not vanish in the region $r_0 < |\tau| < \infty$. Although $V_0 \neq 0$, it is admissible to have the determinant of V_0 equal to zero.

Birkhoff does demonstrate in [1] that the desired reduction of (1) to form (4) with $s = q + 1$ can be brought about by a suitable transformation of type (23), provided:

- (i) that the characteristic roots of A_0 are all distinct and provided
- (ii) that the characteristic roots of J are also all distinct.

As will presently be shown, proviso (ii) is superfluous. Birkhoff indicates this is possibly the case, but gives almost no details when J has multiple elementary divisors.

In the event that proviso (i) is not fulfilled, the author has attempted to show that form (4) with $s = q + 1$ can be reached by a transformation of type (23) when fractional powers of τ do not appear in the asymptotic solutions of (1), but has failed in the more general cases mainly because a suitable modification of Birkhoff's scheme for annulling certain polynomial coefficients has not been found (see [1, pp. 450-452]).

With these preliminary remarks out of the way, we come to

THEOREM III. *If the characteristic roots of A_0 are distinct, there exists a transformation of type (23) which will reduce equation (1) to the canonical form*

$$(25) \quad \tau \frac{dY}{d\tau} = P(\tau)Y,$$

where the elements in the square matrix $P(\tau)$ are polynomials in τ of degree at most $q + 1$.

If $q \leq -2$, the point at $\tau = \infty$ is a regular point for equation (1) and the proof of Theorem III is trivial, for advantage can be taken of the fact that a convergent solution of the form

$$X_f(\tau) = I + \sum_{k=1}^{\infty} X_k/\tau^k$$

exists for equation (1) and therefore a transformation $X = X_f(\tau)Y$ will reduce (1) at once to the form $dY/d\tau = 0$.

Similarly, when $q = -1$, the point $\tau = \infty$ is a regular singular point and equation (1) is known to have a fundamental solution of the form

$$X_f(\tau) = V_1(\tau) \parallel \delta_{ij} \exp\{(\rho_i I_i + J_i) \log \tau\} \parallel,$$

where $V_1(\tau)$ has precisely the structure and properties attributed to matrix $V(\tau)$ in (23) (see [9]). Therefore a transformation $X = V_1(\tau)Y$ will reduce equation (1) to the form

$$\tau \frac{dY}{d\tau} = \parallel \delta_{ij}(\rho_i I_i + J_i) \parallel Y \quad (i, j = 1, \dots, m),$$

making it evident Theorem III is correct when $q = -1$. Here each ρ_i is a constant; the δ_{ij} is a Kronecker delta; the I_i denotes an identity matrix of appropriate order and J_i is likewise a square matrix of the same order as I_i . Moreover every element in J_i is zero except on the first subdiagonal, where 1's or 0's or a mixture of ones and zeros appear.

In the proof of Theorem III, which is about to be given, we can therefore presume $q \geq 0$. Rather than repeat at this juncture in a new notation, pp. 438–451 of Birkhoff's analysis [1], we here summarize the essential facts established by Birkhoff which will enable us to proceed:

(a) If in (1) the roots of A_0 are distinct, equation (1) possesses a fundamental formal series solution

$$(26) \quad \bar{X}(\tau) = U(\tau)H(\tau),$$

where

$$(27) \quad U(\tau) = \sum_{k=0}^{\infty} U_k/\tau^k$$

and

$$(28) \quad H(\tau) = \left\| \delta_{ij} \exp \left\{ \rho_{i,q+1} \log \tau + \sum_{k=0}^q \rho_{i,q-k} \tau^{k+1}/(k+1) \right\} \right\|$$

In (28) the element in the i th row and j th column of H is exhibited. The ρ 's with subscripts are suitably chosen constants and in particular the p_{i0} ($i = 1, \dots, n$), are the n distinct characteristic roots of A_0 . The U_k in (27) are appropriate constant matrices and in particular the determinant of U_0 is not zero.

(b) There exists a transformation of type (23) which will reduce equation (1) to the form

$$(29) \quad \tau \frac{dY}{d\tau} = Q(\tau)Y$$

where the elements in $Q(\tau)$ are polynomials and (29) possesses a formal series solution

$$(30) \quad \tilde{Y}(\tau) = D\{\tau^{\alpha_1}, \dots, \tau^{\alpha_n}\} \Psi(\tau) H(\tau),$$

where the integers $\alpha_1, \dots, \alpha_n$ are so ordered that

$$\alpha_1 \leq \alpha_2 \leq \dots \leq \alpha_n = 0$$

and formally

$$\Psi(\tau) = \sum_{k=0}^{\infty} \Psi_k/\tau^k.$$

The Ψ_k denote constant matrices and the determinant of Ψ_0 is not zero.

(c) Moreover, if we exhibit the individual polynomials in Q by writing

$$Q(\tau) = \| q_{ij}(\tau) \| \quad (i, j = 1, \dots, n),$$

the maximum possible degree of the polynomial $q_{ij}(\tau)$ is $(q + 1 + \alpha_i - \alpha_j)$ and, if this quantity is negative, then $q_{ij}(\tau) \equiv 0$.

If by chance $\alpha_1 = \alpha_n$, the proof of Theorem III is complete. If $\alpha_1 < \alpha_n$, Birkhoff at this point invoked the hypothesis that the roots of J are distinct and began annulling coefficients of certain terms in the polynomials $q_{ij}(\tau)$ involving low powers of τ . To free ourselves from the hypothesis concerning the roots of J we introduce a new procedure for annulling coefficients of certain high powers of τ by means of transformations of type (23), which step by step will bring α_1 closer and closer to α_n , finally making $\alpha_1 = \alpha_n$ and thus proving Theorem III.

We begin by applying to (29) a shearing transformation

$$(31) \quad Y = D\{\tau^{\alpha_1}, \dots, \tau^{\alpha_n}\}W$$

to obtain a new equation

$$(32) \quad \tau \frac{dW}{d\tau} = R(\tau)W$$

with a formal solution $\bar{W} = \Psi(\tau)H(\tau)$, where, if we write $R(\tau) = \|r_{ij}(\tau)\|$, then

$$(33) \quad r_{ij}(\tau) = \sum_{k=\alpha_j-\alpha_i}^{q+1} r_{ijk}\tau^k \quad (i, j = 1, \dots, n),$$

and the r_{ijk} are constants. Here and in future formulas, if a summation becomes meaningless (for instance in (33) the $\alpha_j - \alpha_i > q + 1$), it is understood the summation is to be replaced by a zero.

Since formally

$$R(\tau)/\tau = \frac{d\bar{W}}{d\tau}\bar{W}^{-1} = \frac{d\Psi}{d\tau}\Psi^{-1} + \Psi \frac{dH}{d\tau}H^{-1}\Psi^{-1}$$

and

$$\tau \frac{dH}{d\tau}H^{-1} = \left\| \delta_{ij} \sum_{k=0}^{q+1} \rho_{i,q-k+1}\tau^k \right\|,$$

It is evident that the lead matrix in R , namely

$$(34) \quad R_{q+1} = \|r_{ij,q+1}\| = \Psi_0 \|\delta_{ij}\rho_{i0}\| \Psi_0^{-1}.$$

Therefore the characteristic roots ρ_{i0} of the matrix R_{q+1} are the same as the roots of A_0 and so by hypothesis these roots are all distinct.

There is no loss in generality at this stage, if we assume, as we shall, that in (33) the $r_{ij,q+1} = 0$ if $i > j$ and $r_{ii,q+1} = \rho_{i0}$ for $i = 1, \dots, n$; for, if this were not the case, we would apply a preliminary transformation of the form

$$(35) \quad W = \left\| \begin{array}{cccccc} 1 & 0 & 0 & \cdots & 0 \\ c_{21} & 1 & 0 & & \cdot \\ c_{31} & c_{32} & 1 & & \cdot \\ & \cdots & & \ddots & \cdot \\ c_{n1} & c_{n2} & \cdots & c_{n,n-1} & 1 \end{array} \right\| Z$$

to equation (32) to bring about the desired simplification in R_{q+1} . For details relating to the choice of the constants refer to R. Bellman [10, pp. 21-22]. In particular c_{21}, \dots, c_{n1} are so chosen that

$$(36) \quad (R_{q+1} - \rho_{10}I) \begin{vmatrix} 1 \\ c_{21} \\ \vdots \\ c_{n1} \end{vmatrix} = 0,$$

and ρ_{10} is any one of the roots of A_0 which is not a characteristic root of the $(n-1)$ th ordered matrix formed by erasing the first row and first column in R_{q+1} . There is at least one such root ρ_{10} , for an $(n-1)$ -ordered matrix can not have n distinct roots.

It should be particularly emphasized that transformation (35) will not induce terms into the new equation in Z of lower degree in τ than indicated in (33); i.e., the $r_{ij,q+1}$ in (33) for $i > j$ can be annulled by (35) without changing the lower index of summation ($\alpha_j - \alpha_i$).

Although $\alpha_1, \dots, \alpha_n$ by hypothesis are not all alike, many α 's may be equal. This fact will be taken into account by supposing that

$$(37) \quad \begin{aligned} \alpha_1 &= \alpha_2 = \dots = \alpha_{v_1} = \beta_1; \\ \alpha_{v_1+1} &= \alpha_{v_1+2} = \dots = \alpha_{v_1+v_2} = \beta_2; \\ &\dots \\ \alpha_{v_1+v_2+\dots+v_{m-1}+1} &= \dots = \alpha_{v_1+\dots+v_m} = \alpha_n = \beta_m; \\ \beta_1 &< \beta_2 < \dots < \beta_m \text{ and } v_1 + v_2 + \dots + v_m = n. \end{aligned}$$

With the inequalities (37) in mind, the n by n matrix R is rewritten in a matrix block form, so that it has the structure

$$(38) \quad R(\tau) = \| R_{ij}(\tau) \| \quad (i, j = 1, \dots, m),$$

where the R_{ij} are no longer elements of R , but submatrices, such that R_{ij} has v_i rows and v_j columns. Furthermore

$$(39) \quad R_{ij}(\tau) = \sum_{k=\beta_j-\beta_i}^{q+1} R_{ijk} \tau^k \quad \text{if } i \leq j$$

and

$$(40) \quad R_{ij}(\tau) = \sum_{k=\beta_j-\beta_i}^q R_{ijk} \tau^k \quad \text{if } i > j,$$

where the R_{ijk} are appropriate constant matrices and in particular $R_{ii,q+1}$ has zero elements below the main diagonal and distinct characteristic roots run down the various main diagonals, for $i = 1, \dots, m$.

Indeed we would like to continue the process of annulling elements below the main diagonal using transformations of type (23) and finally obtain an equation of type (32) where (39) holds if $i \leq j$ and

$$(41) \quad R_{ij}(\tau) = \sum_{k=\beta_j-\beta_i}^{q+1-i+j} R_{ijk}\tau^k \quad \text{if } i > j.$$

Once a system (32), (39), (41) has been reached, the proof of Theorem III is essentially complete, for a transformation

$$W = D\{\tau^{m-1}I_1, \tau^{m-2}I_2, \dots, \tau I_{m-1}, I_m\}W_3,$$

where the identity matrix I_i ($i = 1, \dots, m$) has v_i rows, is made and system (32), (39), (41) turns into a new equation of the same type we had in (32), (39), (40), but with one important improvement. The new β_i is equal to the old β_i plus $(m-i)$ ($i = 1, \dots, m$) i.e., the new β 's and new α 's are more nearly alike. The procedure here outlined could then be repeated, for the characteristic roots ρ_{i0} associated with the old and new equations remain unchanged and distinct since $q \geq 0$. Thus in a finite number of steps α_1 would become equal to α_n and the proof of Theorem III would be complete.

Let us therefore backtrack to make sure the desired reduction to (41) can be carried out. Already (41) is satisfied if $(i-j) = 1$; see (40). To carry out an induction suppose that transformations of type (23) have been carried out and (39) holds with all elements in $R_{ii, q+1}$ zero below the main diagonal and the distinct characteristic roots are still running down the main diagonals, for $i = 1, \dots, m$. Also suppose that

$$(42) \quad R_{ij}(\tau) = \sum_{k=\beta_j-\beta_i}^{q+1-i+j} R_{ijk}\tau^k$$

when $i-j = 1, \dots, \mu < m-1$ and that

$$(43) \quad R_{ij}(\tau) = \sum_{k=\beta_j-\beta_i}^{q+1-\mu} R_{ijk}\tau^k$$

for $i-j = \mu+1, \dots, m-1$.

To push forward the reduction, replace the block structure of R indicated in (38) by the cruder block structure

$$(44) \quad R = \|\mathcal{R}_{ij}(\tau)\| \quad (i, j = 1, 2, 3),$$

where the square submatrices \mathcal{R}_{11} , \mathcal{R}_{22} , and \mathcal{R}_{33} have the respective orders v_1 , $(v_2 + v_3 + \dots + v_{\mu+1})$ and $(v_{\mu+2} + \dots + v_m)$. Also \mathcal{R}_{12} has v_1 rows and $(v_2 + \dots + v_{\mu+1})$ columns, etc.

A transformation

$$W = \begin{vmatrix} I_1 & 0 & 0 \\ 0 & I_2 & 0 \\ C\tau^{-\mu} & 0 & I_3 \end{vmatrix} W_4,$$

where the identity matrix I_i ($i = 1, 2, 3$) is of the same order as matrix \mathcal{R}_{ii} in (44), is applied to (32) to obtain a new equation

$$(45) \quad \frac{dW_4}{d\tau} = \frac{\|S_{ij}(\tau)\|}{\tau} W_4 \quad (i, j = 1, 2, 3).$$

A computation will show that

$$\begin{aligned} S_{ij}(\tau) &= \mathcal{R}_{ij}(\tau) & (i = 1, 2; \quad j = 2, 3); \\ S_{i1}(\tau) &= \mathcal{R}_{i1}(\tau) + \mathcal{R}_{i3}(\tau)C/\tau^\mu & (i = 1, 2); \\ (46) \quad S_{3j}(\tau) &= \mathcal{R}_{3j}(\tau) - C\mathcal{R}_{1j}(\tau)/\tau^\mu & (j = 2, 3); \\ S_{31}(\tau) &= \mathcal{R}_{31}(\tau) + \{\mathcal{R}_{33}(\tau)C - C\mathcal{R}_{11}(\tau)\}/\tau^\mu - C\mathcal{R}_{13}(\tau)C/\tau^{2\mu} + \mu C/\tau^\mu. \end{aligned}$$

The C can and will be so chosen that in $S_{31}(\tau)$ the terms to the highest power of τ (i.e., τ to the power $q + 1 - \mu$) cancel out.

This possibility is based on

LEMMA I. *A matrix equation of the form*

$$(47) \quad AX - XB = C$$

for given matrices A, B, C always has a unique solution X if A and B have no common characteristic root.

For a proof of this lemma, see reference [11]. It is understood in (47) that if the square matrix A has m rows and the square matrix B has n rows, the matrices X and C each have m rows and n columns.

The various equations in (46) can then be used to verify that the new equation (45) has precisely the same form as the old equation (32), where, after changing S 's into R 's, the new block matrices $R_{ij}(\tau)$ satisfy (39), (42) and (43) and furthermore because of the judicious choice of C the new

$$(48) \quad R_{j\mu, q+1-\mu} = 0 \quad \text{for } i = \mu + 2, \mu + 3, \dots, m.$$

It is particularly important to check here the fact, as one can easily do using (46), that the upper indices of summation on the \sum 's in the formulas describing the new R_{ij} have not become larger than here stated and that the lower indices of summation are not below the values here specified and that the characteristic roots ρ_{i0} have not been disturbed.

To complete the reduction, let it be assumed that at this stage in the analysis we are dealing with an equation (32) where the R_{ij} satisfy not only (39), (42) and (43), but that when $j = 1, 2, \dots, \eta$, where η is some particular integer on the range $1, \dots, m - \mu - 2$, then also

$$(49) \quad R_{ij, q+1-\mu} = 0 \quad \text{for } i = \mu + j + 1, \dots, m.$$

To carry the annulling one step further and get an equation of type (32) where not only (39), (42), (43) and (49) are satisfied, but also where

$$(50) \quad R_{i, \eta+1, q+1-\mu} = 0 \quad \text{for } i = \mu + \eta + 2, \dots, m,$$

first split the matrix R in (32) into the cruder block form

$$(51) \quad R = \parallel \mathfrak{R}_{ij}(\tau) \parallel \quad (i, j = 1, \dots, 4),$$

where the square matrix \mathfrak{R}_{11} has $(v_1 + v_2 + \dots + v_\eta)$ rows; the square matrix \mathfrak{R}_{22} has $v_{\eta+1}$ rows, \mathfrak{R}_{33} has $(v_{\eta+2} + v_{\eta+3} + \dots + v_{\mu+\eta+1})$ rows and an equal number of columns and \mathfrak{R}_{44} has $(v_{\mu+\eta+2} + \dots + v_m)$ rows and a like number of columns, etc.

Then make a transformation of the form

$$(52) \quad W = \left\| \begin{array}{cccc} I_1 & 0 & 0 & 0 \\ 0 & I_2 & 0 & 0 \\ 0 & 0 & I_3 & 0 \\ 0 & C\tau^{-\mu} & 0 & I_4 \end{array} \right\| W_5,$$

where the identity matrix I_i ($i = 1, \dots, 4$) has the same number of rows and columns as the \mathfrak{R}_{ii} in (51). This converts (32) into a new equation of the form

$$(53) \quad \tau \frac{dW_5}{d\tau} = \parallel S_{ij} \parallel W_5 \quad (i, j = 1, \dots, 4).$$

Formulas analogous to (46) are easily found expressing the new S_{ij} matrices in (53) in terms of the \mathfrak{R}_{ij} in (51) and again it is not difficult to verify that the constant matrix C in (52) can and will be so chosen that not only does the new equation (53) have structure (32), but also that, when in this equation the S 's are replaced by R 's and the resultant finer R_{ij} block structure is considered, it will satisfy all the desired relations (39), (42), (43), (48), (49), and (50).

The procedure thus outlined, with η increased a unit each time, can therefore be repeated, always using transformations of type (23) and each time modifying (32) more and more until not only (39), (42), (43) and (48) are satisfied, but also (49) with $\eta = m - \mu - 1$. At this stage one notices that not only are the relations (32), (39), (42), (43) satisfied, but also the same relations are now satisfied when in these formulas μ is replaced by $\mu + 1$.

The process then moves forward with increasing μ until μ has become equal

to $(m-1)$ and at this point the reduction is complete and Theorem III is demonstrated as previously explained, since the R_{ij} in (32) at this juncture satisfy (39) and (41).

4. A regular singular point. We now turn to the special case $q = -1$ and present

THEOREM IV. *If $q = -1$ in equation (1), then there exists a transformation of type (3) which will reduce (1) to a canonical form of type (4) where $s = 1$.*

To prove Theorem IV it is convenient to shift the regular singular point at $\tau = \infty$ to the origin $t = 0$ by setting $\tau = 1/t$ and absorb a minus sign in the constant coefficients so that (1) takes the revised form

$$(58) \quad t \frac{dX}{dt} = \sum_{v=0}^{\infty} A_v t^v X.$$

At first we seek a transformation of the form

$$(59) \quad X = T(t)Y,$$

where T possesses a convergent series representation

$$(60) \quad T = \sum_{v=0}^{\infty} T_v t^v, \quad \text{where } T_0 = I,$$

which for some choice of the integer w , will convert (58) into a new equation

$$(61) \quad t \frac{dY}{dt} = \sum_{v=0}^w A_v t^v Y,$$

where A_0, A_1, \dots, A_w in (58) are the same matrices as in (61). Any T of the form (60) will bring about this desired reduction of (58) to (61), provided T is a solution of the differential equation

$$(62) \quad t \frac{dT}{dt} = \sum_{v=0}^{\infty} A_v t^v T - T \sum_{v=0}^w A_v t^v.$$

Proceeding temporarily in a formal fashion, let us substitute (60) into (62) and equate coefficients of like powers of T to obtain the recurrence relations:

$$T_1 A_0 - (A_0 - I) T_1 = 0;$$

$$(63) \quad T_r A_0 - (A_0 - rI) T_r = \sum_{v=1}^{r-1} A_v T_{r-v} - \sum_{v=1}^{r-1} T_{r-v} A_v \quad (r = 2, \dots, w);$$

$$T_r A_0 - (A_0 - rI) T_r = \sum_{v=1}^{r-1} A_v T_{r-v} - \sum_{v=1}^w T_{r-v} A_v + A_r \quad (r > w).$$

It is evident that, for all sufficiently large values of r , the characteristic roots of A_0 will be distinct from the roots of $(A_0 - rI)$. To be more specific, suppose that when $r \geq r_0 > 0$ no root of A_0 coincides with a root of $(A_0 - rI)$. Then take $w = r_0 - 1$ and this choice will guarantee there is a sequence of constant matrices T_1, T_2, \dots satisfying system (63), for one takes

$$(64) \quad T_1 = T_2 = \dots = T_{r_0-1} = 0$$

and then $T_{r_0}, T_{r_0+1}, \dots$ are uniquely determined by the recurrence relations (63). Thus at least a formal solution (60) has been found for equation (62).

In writing (63) it has been tacitly assumed that $r_0 - 1 = w \geq 2$. If by chance $r_0 - 1 = w = 1$, the second equation in (63) is to be omitted; set $T_1 = 0$ and T_2, T_3, \dots are then all uniquely determined in turn. When $r_0 - 1 = w = 0$, omit both the first and second equations in (63) and T_1, T_2, \dots are all determined in turn by the third equation in (63). Hence in any case a formal series solution (60) for equation (62) has been found.

Equation (62) does not have form (58), but can easily be made to take on a structure analogous to (58) by introducing as the unknown vector

$$V = \begin{pmatrix} T^{(1)} \\ T^{(2)} \\ \vdots \\ T^{(n)} \end{pmatrix},$$

where $T^{(j)}$ is the j th column in matrix T . It is then apparent that this vector V satisfies a differential equation of the form

$$(65) \quad t \frac{dV}{dt} = \sum_{v=0}^{\infty} B_v t^v V,$$

where the indicated series in (65) converges for sufficiently small values of t . The B_v could be expressed in terms of the elements in the A_v matrices, but these relations are not needed for present purposes. The important feature is that since (58) has a formal matrix series solution of type (60), (65) also has a formal vector solution of the form

$$(66) \quad V = \sum_{k=0}^{\infty} V_k t^k.$$

But we know any formal vector solution of type (66) of equation (65) must necessarily converge for sufficiently small $|t|$. This in turn implies series (60) converges and we have as a result

LEMMA II. *If for some positive integer r , no characteristic root of matrix A_0 in (58) is equal to a root of $(A_0 - rI)$ and this is also true for all larger*

integral values of r , then there exists a convergent transformation of type (59) such that (59) will reduce (58) to the form

$$(67) \quad t \frac{dY}{dt} = \sum_{v=0}^{r-1} A_v t^v Y,$$

where the A_0, \dots, A_{r-1} in (58) and in (67) are the same respective matrices.

Returning to a general equation of type (58), there is no loss of generality in assuming, as we shall, that A_0 is in the Jordan classical canonical form. Matrix A_0 can then be written in a cruder form

$$A_0 = \|\delta_{ij} A_{ii0}\| \quad (i, j = 1, \dots, m),$$

where the characteristic roots have been so arranged that the differences between the roots of A_{ii0} and A_{jj0} are not integers if $i \neq j$. The roots of any particular matrix A_{ii0} are either all alike or they may differ one from the other by integers.

Let the maximum difference between two roots of A_{ii0} be denoted by $(r_i - 1)$ and let $r = \max\{r_1, \dots, r_m\}$.

In reference [9] it is shown that a convergent transformation of type (59) exists that will convert (68) into a new equation of the form

$$(68) \quad t \frac{dY}{dt} = \left\| \delta_{ij} \sum_{v=0}^{r-1} A_{iiv} t^v + \sum_{v=r}^{\infty} A_{ijv} t^v \right\| Y.$$

But according to Lemma II an appropriate convergent transformation of type (59) will obliterate the second summation in (68) without changing the terms in the first summation. The resulting modified form of (68) with the second summation missing will be equivalent to m separate matrix equations because of the diagonal block structure of the modified form of (68).

In short there is no loss in generality in confining the analysis to equations of type (58) where A_0 is not only in Jordan form, but also the roots of A_0 are alike or differ by integers. It has been shown in reference [9] that with such a lead matrix A_0 there will be a transformation of type (60) which will further reduce the equation under consideration to the form

$$(69) \quad t \frac{dZ}{dt} = \left\| \delta_{ij} (\rho_i I_i + J_i) + K_{ij}(t) + \sum_{v=k_i+1}^{\infty} B_{ijv} t^v \right\| Z,$$

where

$$(70) \quad \rho_i = \rho_m + k_i \quad \text{for } i = 1, \dots, m$$

and the integers k_i are so ordered that

$$(71) \quad k_1 > k_2 > \dots > k_m = 0.$$

Also

$$(72) \quad K_{ij}(t) = 0 \quad \text{if } i \geq j;$$

$$(73) \quad K_{ij}(t) = B_{ij}t^{k_i - k_j} \quad \text{if } i < j,$$

and all the B_{ij} 's are appropriate constant matrices. The infinite series in (69) converges for sufficiently small $|t|$.

It will now be shown, that, given an equation of type (69) with the roots ρ_i satisfying (70) and (71), there will be associated with (69) another equation

$$(74) \quad t \frac{dX}{dt} = \|\delta_{ij}(\rho_i I_i + J_i) + t C_{ij}\| X \quad (i = 1, \dots, m),$$

where

$$(75) \quad C_{ij} = 0 \quad \text{if } i > j; \quad C_{11} = 0; \quad \text{and } C_{ii} = \gamma_i I_i,$$

and a transformation of type (59), namely

$$(76) \quad X = \|\delta_{ij} I_i + Q_{ij}(t)\| Y,$$

where

$$Q_{ij}(t) = 0 \quad \text{if } i > j; \quad Q_{11}(t) = 0$$

and

$$(77) \quad Q_{ij}(t) = \sum_{v=1}^{k_1} Q_{ijv} t^v \quad \text{if } i \leq j,$$

such that transformation (76) will reduce (74) to an equation of the form

$$(78) \quad t \frac{dY}{dt} = \|\delta_{ij}(\rho_i I_i + J_i) + K_{ij}(t) + \sum_{v=k_1+1}^{\infty} C_{ijv} t^v\| Y,$$

where the infinite series in this equation converges for sufficiently small values $|t|$; $C_{11v} = 0$ and $C_{ijv} = 0$ if $i > j$ and $v = k_1 + 1, k_1 + 2, \dots$. In (75) the γ_i is a nonzero scalar parameter, which eventually will be set equal to i .

By virtue of Lemma II the B_{ijv} in (69) and the C_{ijv} in (78) can be eliminated without affecting the other terms in these two equations by means of transformations of type (59). In this way the two equations (69) and (78) can be modified and made identical. Since inverses and products of matrices of type (60) are also of type (60), we see (69) can be reduced to an equation of form (74) by a transformation of type (59) and thus Theorem IV will be demonstrated, provided the aforementioned reduction of (75) to (78) can be carried out. To examine this possibility, substitute (76) into (74) and compute the resulting coefficient matrix of Y in the new equation thus obtained. Then equate this coefficient matrix to the coefficient matrix of Y indicated in (78). Having done this, multiply on the left

both sides of the resulting equation by the matrix $\|\delta_{ij}I_i + Q_{ij}(t)\|$ and compare coefficients of like powers of t on both sides of the resulting equation.

It will be found, first of all, one must select the Q_{iiv} 's so that

$$(79) \quad Q_{ii1} + Q_{ii1}J_i - J_iQ_{ii1} = C_{ii} \quad (i = 2, \dots, m)$$

and

$$(80) \quad rQ_{iir} + Q_{iir}J_i - J_iQ_{iir} = C_{ii}Q_{ii,r-1} \quad (i = 2, \dots, m; r = 2, \dots, k),$$

where in formulas (80) through (88) a k is being written in place of k_1 .

Also for $i = 1, \dots, m; j = 2, \dots, m$ and $i < j$ we must be able to select the constant matrices C_{ij} and Q_{ijv} so that in the equation

$$(81) \quad \begin{aligned} & k_i \sum_{v=1}^k Q_{ijv}t^v + J_i \sum_{v=1}^k Q_{ijv}t^v + \sum_{\eta=i}^j \sum_{v=1}^k C_{i\eta}Q_{\eta jv}t^{v+1} + tC_{ij} \\ & = k_j \sum_{v=1}^k Q_{ijv}t^v + \sum_{v=1}^k Q_{ijv}t^v J_j + B_{ij}t^{k_i-k_j} + \sum_{v=1}^k vQ_{ijv}t^v \\ & \quad + \sum_{\eta=i}^{j-1} \sum_{v=1}^k Q_{i\eta v}B_{\eta j}t^{v+k_\eta-k_j} + \text{terms of degree } (k_1+1) \text{ and higher in } t, \end{aligned}$$

the coefficients of like powers of t on both sides of (81) are the same for t, t^2, \dots, t^k . If the C_{ij} and Q_{ijv} can be so selected as to meet these conditions the desired reduction can be carried out.

The presence in these formulas of the J_i 's is a complicating feature. An artifice will be introduced that will in effect make the influence of the J_i matrices negligible.

It will be assumed, without loss of generality, that in every matrix J_i ($i=1, \dots, m$) in (69) and in all our subsequent equations the 1's running down the first subdiagonal of each J_i have all been replaced by ε 's, where ε is an arbitrarily small positive quantity. This assumption in effect means that a preliminary transformation

$$(82) \quad Z = D(\varepsilon^{n-1}, \varepsilon^{n-2}, \dots, \varepsilon, 1)W$$

has been applied to (69). Since we want to keep ε a free variable, one should note that transformation (82) the roots ρ_i in (69) remain unchanged and independent of ε , while the matrices B_{ijv} become continuous functions of ε .

With the new meaning of the symbol J_i in mind, turn to equations (79) and (80) and note that since $C_{ii} = \gamma_i I_i$, Q_{ii1} is uniquely determined by (79) by virtue of Lemma I. Moreover

$$(83) \quad \lim_{\varepsilon \rightarrow 0} Q_{ii1} = \gamma_i I_i \quad (i = 2, \dots, m).$$

Likewise in turn Q_{ii2}, \dots, Q_{iik} are computed from (80) and

$$(84) \quad \lim_{\varepsilon \rightarrow 0} Q_{iir} = \gamma_i^r I_i / r! \quad (r = 1, \dots, k; i = 2, \dots, m).$$

Temporarily assume $k - k_2 \neq 1$ and set

$$(85) \quad Q_{12, k-k_2} = 0.$$

Then in (81) let $i = 1$ and $j = 2$ and equate coefficients of like powers of t^{k-k_2} to find that

$$(86) \quad C_{12} Q_{22, k-k_2-1} = B_{12}.$$

By choosing ε small enough, it is clear from (83) and (84) there is no danger of $Q_{22, k-k_2-1}$ being a singular matrix. Therefore C_{12} is uniquely determined by (86).

Keeping $i = 1$ and $j = 2$ in (81) next equate coefficients of t and one finds

$$(87) \quad kQ_{121} + J_1 Q_{121} + C_{12} = Q_{121} J_2 + Q_{121} + k_2 Q_{121}$$

and (87) uniquely determines Q_{121} . Likewise if $r \geq 2$, $r \leq k$, and $r \neq k - k_2$,

$$(88) \quad kQ_{12r} + J_1 Q_{12r} + C_{12} Q_{22, r-1} = Q_{12r} J_r + rQ_{12r} + k_2 Q_{12r}$$

and from (88) the unknowns $Q_{122}, Q_{123}, \dots, Q_{12k}$ are computed in turn.

In the special case where $k - k_2 = 1$, take

$$(89) \quad Q_{121} = 0 \quad \text{and} \quad C_{12} = B_{12}$$

and again the remaining matrices Q_{122}, \dots, Q_{12k} are computed in turn from equation (88).

We are ready to set $i = 2$ and $j = 3$ in (81) and equate coefficients of like powers of t . One finds, if $k_2 - k_3 \neq 1$, that

$$(90) \quad k_2 Q_{231} + J_2 Q_{231} + C_{23} = k_3 Q_{231} + Q_{231} J_3 + Q_{231};$$

and

$$(91) \quad k_2 Q_{23r} + J_2 Q_{23r} + C_{22} Q_{23, r-1} + C_{23} Q_{33, r-1} = k_3 Q_{23r} + Q_{23r} J_3 + rQ_{23r}$$

for $r = 2, \dots, k_2 - k_3 - 1$. Setting

$$Q_{23, k_2 - k_3} = 0,$$

one also finds

$$(92) \quad C_{22} Q_{23, k_2 - k_3 - 1} + C_{23} Q_{33, k_2 - k_3 - 1} = B_{23}.$$

The situation has become more involved. First, one solves (90) for Q_{231} and the elements in Q_{231} will then be given as certain linear combinations of the elements in C_{23} . The elements of C_{23} are as yet undetermined. Then, one solves (91) in turn for Q_{232} , then Q_{233} , and so on finally obtaining the elements in $Q_{23, k_2 - k_3 - 1}$ expressed as linear combinations of the elements in C_{23} . This brings

us then to the crucial equation (92), where we must be sure one can solve for the elements in C_{23} .

To see there certainly will be a unique solution of (92) when ε is chosen small enough, consider the equation

$$(93) \quad Q(1) = N(1) + C/(1-h),$$

where h is any positive integer greater than one, and the equations

$$(94) \quad Q(r) = \frac{\gamma Q(r-1)}{r-h} - \frac{\mu^{r-1}C}{(r-1)!(h-r)} + N(r),$$

where $r = 2, \dots, h-1$; γ and μ are positive constants; $N(1)$ is a known matrix independent of C and $N(r)$ is some known matrix function of r which is also independent of C .

An evaluation of

$$(95) \quad \mathcal{R} = \gamma Q(h-1) + C\mu^{h-1}/(h-1)!$$

is needed. To this end compute $Q(2), Q(3), \dots, Q(h-1)$ in succession and substitute the result found for $Q(h-1)$ in (95).

For instance one finds

$$(96) \quad Q(2) = \{N(2) + \gamma N(1)/(2-h)\} + C\{\gamma + \mu(1-h)\}/(2-h)(1-h),$$

and without difficulty one obtains

$$(97) \quad \mathcal{R} = \Omega + C(\mu - \gamma)^{h-1}/(h-1)!,$$

where Ω denotes a known matrix which is independent of C . For present purposes nothing more need be known about the Ω matrix.

As the procedure here described is carried forward, we will need to use (93), (94) and (97) when

$$h = k_i - k_j; C = C_{ij}; \gamma = \gamma_i; \mu = \gamma_j$$

and in each instance one must be sure

$$(\mu - \gamma) = (\gamma_j - \gamma_i) \neq 0 \quad \text{when } i < j;$$

so we set $\gamma_i = i$ for $i = 2, \dots, m$.

Let us return to system (90)–(91) and set $\varepsilon = 0$, (i.e., assume by chance the J 's are all identically zero). System (90)–(91) then turns into system (93)–(94) with

$$Q(r) = Q_{23r}, \quad C = C_{23}, \quad \gamma = \gamma_2, \quad \mu = \gamma_3$$

and $h = k_2 - k_3$. In this instance all N 's and Ω are zero. The left member of (92) becomes the \mathcal{R} given in (97) and it is at once clear one can solve for $C = C_{23}$, provided $\varepsilon = 0$. But the continuity of the elements in our matrices as functions of

ε guarantees that (92) uniquely determines C_{23} even if ε is not zero, provided the ε is chosen small enough.

Thus with C_{23} determined, the $Q_{231}, Q_{232}, \dots, Q_{23, k_2 - k_3 - 1}$ likewise become determined and there is no further difficulty in computing $Q_{23, k_2 - k_3 + 1}, \dots, Q_{23k}$ in turn by equating coefficients of the respective powers of $t^{k_2 - k_3 + 1}, \dots, t^k$ in (81) with $i = 2$ and $j = 3$.

In the special case when $k_2 - k_3 = 1$, set $Q_{231} = 0$ and then $C_{23}, Q_{232}, \dots, Q_{23k}$ can easily be computed, one after the other in the order listed.

The process of evaluating the C_{ij} 's and Q_{ijv} 's then moves forward step by step. Next comes the evaluation of C_{34} and the Q_{34v} for $v = 1, \dots, k$. One moves in this way down the first superdiagonal until all the C_{ij} and Q_{ijv} for $i + 1 = j$ have been evaluated; then one begins on the second superdiagonal, where $i + 2 = j$ and evaluates the C_{13} and the Q_{13v} 's and so moves on down the second superdiagonal to the right. Then comes the third superdiagonal and we work from top down to the right, etc.

As we move ahead in each instance set

$$Q_{ij, k_i - k_j} = 0 \quad \text{for } i = 2, \dots, m; j = 3, \dots, m; i < j.$$

Since the computations at each stage are analogous to those already described save that nonzero N 's and Ω 's may appear, no details need be given.

Thus all the C_{ij} 's and Q_{ijv} 's can be evaluated in turn and Theorem IV is proved.

5. Summary. In §2 the source of G. D. Birkhoff's error has been pointed out. It is found that the s in (4) is $(q + 1)$ if and only if the E_i 's, ($i = \alpha + 1, \dots, \beta$), are not present in the monodromic matrix (7) or, if the E_i 's are present, in the corresponding matrices K_i the α_{ij} 's are so ordered that (22) is satisfied.

Birkhoff's original reduction theorem has been extended in §3 by dropping one of his two hypotheses.

The special case $q = -1$ is treated in full detail in §4 and the lowest possible value for s is obtained.

When $q \geq 0$, the author would have liked to have determined the minimal value of s in (4) that can be obtained by using transformations of not only types (3) and (23), but also for a more general type of nonsingular transformation, where in (24) it would be admissible to replace τ by a fractional power of τ . However to-date he has not been able to obtain such general results.

REFERENCES

1. G. D. Birkhoff, *Singular points of ordinary linear differential equations*, Trans. Amer. Math. Soc. **10** (1909), 446.
2. ———, *Equivalent singular points of ordinary linear differential equations*, Math. Ann. **74** (1913), 136.
3. F. R. Gantmacher, *The theory of matrices*, Vol. 2, Chelsea, New York, 1959; English transl. of Gosudarstv. Izdat. Tech.-Teor. Lit., Moscow, 1953, p. 491 (Russian).

4. P. Masani, *On a result of G. D. Birkhoff on linear differential systems*, Proc. Amer. Math. Soc. **10** (1959), 696–698.
5. G. D. Birkhoff, *A theorem on matrices of analytic functions*, Math. Ann. **74** (1913), 122–133.
6. ———, *On a simple type of irregular singular point*, Trans. Amer. Math. Soc. **14** (1913), 463.
7. D. Hilbert, *Grundzüge einer allgemeinen Theorie der linearen Integralgleichungen*, Göttinger Nachr. (1905), 307–338.
8. J. Plemelj, *Riemannsche Funktionenscharen mit gegebener Monodromiegruppe*, Monatsh. Math. Phys. **19** (1908), 211–246.
9. H. L. Turrittin, *Convergent solutions of ordinary linear homogeneous differential equations in the neighborhood of an irregular singular point*, Acta Math. **93** (1955), 27–66.
10. R. Bellman, *Stability theory of differential equations*, McGraw-Hill, New York, 1953.
11. C. E. Cullis, *Matrices and determinoids*. I, Cambridge Univ. Press, London, 1925, Vol. 3, pp. 419–421 and 441.

UNIVERSITY OF MINNESOTA,
MINNEAPOLIS, MINNESOTA