MATRIX PRODUCTS OF MATRIX POWERS

R. F. CLIPPINGER

1. Introduction. Let m n-by-n matrices, A_k, of complex constants, a_{ijk} $(i, j=1, 2, \cdots, n; k=1, 2, \cdots, m)$, be given. We shall denote by $\mathcal{L}$ the set of all matrices,

$$A(t) = \sum_{i=1}^{m} \rho_i(t)A_i,$$

where $\rho_i(t)$ $(i=1, 2, \cdots, m)$ are arbitrary, non-negative, summable functions of the real variable t on the interval T, $a \leqq t \leqq b$. We shall call $\mathfrak{J}$, $\mathfrak{S}$, or $\mathfrak{X}$ the subsets of $\mathcal{L}$ obtained by restricting the functions $\rho_i(t)$ to polynomial functions, step functions, or step functions which are all zero except one. Since, in each case, the elements of $A(t)$ are summable functions of t on T, it follows that, on T, there exists a unique, absolutely continuous matrix solution,[1] $Y(t)$, of the linear, matrix differential equation and initial condition:

$$dY(t)/dt = Y(t)A(t), \qquad Y(a) = E, \tag{1.1}$$

where E is the n-by-n unit matrix. We shall denote by λ, ι, σ or ξ the set of matrices, $Y(t)$, which are particular values of solutions of (1.1), where $A(t)$ is an arbitrary matrix of $\mathcal{L}$, $\mathfrak{J}$, $\mathfrak{S}$, or $\mathfrak{X}$, respectively, and t is on T.

If A is a matrix with elements a_{ij}, let the absolute value of A and the exponential and natural logarithm of A be defined[2] by the equations:

$$|A| = \left[\sum_{i,j=1}^{n} |a_{ij}|^2\right]^{1/2},$$

$$\exp A = \sum_{i=0}^{\infty} A^i/i!$$

$$\log A = \sum_{i=1}^{\infty} (-1)^{i-1}(A - E)^i/i, \quad \text{if} \quad |A - E| < 1.$$

Presented to the Society, April 18, 1942; received by the editors April 4, 1943, and, in revised form, August 18, 1943. The author wishes to thank Professor G. D. Birkhoff for suggestions which led to this paper.

[1] See W. M. Whyburn, *On the fundamental existence theorems for differential systems*. Ann. of Math. (2) vol. 30 (1928–29) p. 31. We observe that equations (1.1) are equivalent to a system of $2n$ real, linear, first order differential equations satisfying all the hypotheses of this theorem.

[2] See J. v. Neumann, *Über die analytischen Eigenschaften von Gruppen linearer Transformationen und ihrer Darstellungen*. Math. Zeit. vol. 30 (1929) pp. 6, 7.

J. v. Neumann[3] has shown that if j is an integer and if $|A-E|<1$, then

$$A^j = \exp\,[j \log A]. \tag{1.2}$$

Hence, for any α, we define the α power function of A by the equation

$$A^\alpha = \exp\,(\alpha \log A), \quad \text{if} \quad |A - E| < 1.$$

If $|\exp A_i - E| < 1$ $(i=1, 2, \cdots, m)$, we define μ as the set of matrix products of matrix powers,

$$\prod_{j=1}^{J} \prod_{i=1}^{m} (\exp A_i)^{\alpha_{ij}},$$

where the α_{ij} are arbitrary non-negative numbers.

Let us identify an n-by-n matrix, B, of complex numbers with the point in $2n^2$-Euclidean space, whose coordinates are the real and imaginary parts of the elements of B. The distance between two points B_1 and B_2 may be defined as $|B_1-B_2|$. A set, β, of matrices, B, is then also a point set whose closure we denote by $\bar{\beta}$.

It is the purpose of this paper to show that the sets $\bar{\lambda}$, $\bar{\iota}$, $\bar{\sigma}$, $\bar{\xi}$ and, if it exists, $\bar{\mu}$, are identical.

2. Principal theorems.

THEOREM 2.1. *The sets $\bar{\lambda}$ and $\bar{\sigma}$, defined above, are identical.*

Since any step function on T is summable on T, it follows that $\sigma \subset \lambda$. Suppose $A_{\mathrm{L}}(t)$ is a matrix of class $\mathcal{L}$ with coefficients $\rho_i(t)$. Then, for all positive δ, there exists a matrix, A_{S}, of class $\mathfrak{S}$ with coefficients $r_i(t)$ such that

$$\sum_{k=1}^{m} \int_a^b |\rho_k(t) - r_k(t)|\,|a_{ijk}|\,dt < \delta \qquad (i, j = 1, 2, \cdots, n).$$

Let the corresponding solutions of (1.1) be $Y_{\mathrm{L}}(t)$ and $Y_{\mathrm{S}}(t)$. Since $Y(t)$ is a uniformly continuous functional[4] of $A(t)$, it follows that, given any positive number ϵ, δ may be so chosen that

$$Y_{\mathrm{L}}(t) - Y_{\mathrm{S}}(t) \ll \epsilon;$$

that is, the absolute value of each element of the matrix on the left is less than ϵ. Hence $\lambda \subset \bar{\sigma}$.

---

[3] Loc. cit. pp. 8, 12.

[4] By this we mean that the elements of $Y(t)$ are uniformly continuous functionals of the elements of $A(t)$. See W. M. Whyburn, *Functional properties of the solutions of differential systems*. Trans. Amer. Math. Soc. vol. 32 (1930) p. 508.

THEOREM 2.2. *The sets* $\bar{\sigma}$ *and* ξ *are identical.*

Clearly $\xi \subset \sigma$.

J. v. Neumann[5] has shown that if $|A|, |B| < (1/2) \log (3/2)$, then

$$\log [\exp A \exp B] = A + B + O(|A||B|).$$

This equation, by induction, leads to the generalized equation:

$$\log \prod_{i=1}^{m} \exp A_i = \sum_{i=1}^{m} A_i + \sum_{i,j=1}^{m} O(|A_i||A_j|), \tag{2.1}$$

if $|A_i| < \delta(m)$, where it suffices to take $\delta(m) < [\log (3/2)]/2n(m-1)$. To this we may add the equation,

$$\exp (A + B) = \exp A + O(|B|) \tag{2.2}$$

which follows immediately from the definition of $\exp (A+B)$.

LEMMA 2.1. *If A is a matrix of constants, the matrix,*

$$Y(t) = Y_0 \exp [(t - a)A],$$

is the solution of the linear, matrix differential equation and initial condition

$$dY(t)/dt = Y(t)A, \qquad Y(a) = Y_0.$$

The series $\sum_{j=0}^{\infty} (t-a)^j A^j / j$ is uniformly convergent[6] on any interval $|t-a| < N$, hence the lemma may be established by term-by-term differentiation.

LEMMA 2.2. *If $A(t)$ is a matrix of summable functions, and if $A(t) \ll (M)$ on T, the solution, $Y(t)$, of equation* (1.1) *satisfies the inequality*

$$Y(t) - E \ll (1/n[\exp Mn(t - a) - 1]) \text{ on } T.$$

Slight modifications of the proof of the existence theorem given by G. D. Birkhoff and R. E. Langer[7] yield this lemma.

LEMMA 2.3. *If $B_1, B_2, \cdots, B_m$ are square matrices,*

$$\lim_{j\to\infty} \left[\prod_{i=1}^{m} \exp (B_i/j) \right]^j = \exp \sum_{i=1}^{m} B_i.$$

---

[5] Loc. cit. pp. 13–15.

[6] See J. v. Neumann, loc. cit. p. 7.

[7] *The boundary problems and developments associated with a system of ordinary linear differential equations of the first order.* Proceedings of the American Academy of Arts and Sciences vol. 58 (1922–1923) pp. 59–63.

Let $P_j=\prod_{i=1}^{m}\exp(B_i/j)$. By Lemma 2.1, P_j is the solution for $t=1$ of the differential equation

$$dY(t)/dt = Y(t)\sum_{i=1}^{m}\rho_i(t)B_i, \qquad Y(0) = E,$$

where the $\rho_i(t)$ are all zero except on the interval $(i-1)/m \leqq t \leqq i/m$ where $\rho_i(t)=m/j$ $(i=1, 2, \cdots, m)$. Let B be an upper bound to the absolute values of the elements of mB_i. Then $\sum_{i=1}^{m}\rho_i(t)B_i \ll (B/j)$. By Lemma 2.2, j may be chosen so large that $|P_j-E|<1$ and $|B_i/j| < [\log(3/2)]/2n(m-1)$. Hence by (1.2)

$$P_j = \exp\left[j\log\prod_{i=1}^{m}\exp(B_i/j)\right].$$

From (2.1) it follows that

$$j\log\prod_{i=1}^{m}\exp(B_i/j) = \sum_{i=1}^{m}B_i + O(1/j);$$

and, finally, (2.2) implies that

$$P_j = \exp\sum_{i=1}^{m}B_i + O(1/j).$$

This establishes Lemma 2.3.

LEMMA 2.4. *If each of K n-by-n matrices, C_k, is the limit of a sequence, $\{C_{kj}\}$, of matrices, then*

$$\lim_{j\to\infty}\prod_{k=1}^{K}C_{kj} = \prod_{k=1}^{K}C_k.$$

This lemma may be established easily by induction.

To prove that $\sigma\subset\bar{\xi}$, let Y_S be a matrix of σ; then Y_S is a product of a finite number of matrices of the form

$$C_k = \exp\sum_{i=1}^{m}\alpha_{ik}A_i,$$

and by Lemma 2.3, each C_k is a limit of matrices

$$C_{kj} = \left[\prod_{i=1}^{m}\exp(\alpha_{ik}A_i/j)\right]^j.$$

Hence, by Lemma 2.4, Y_S is a limit of matrices of ξ and therefore Y_S is a member of $\bar{\xi}$.

COROLLARY TO THEOREM 2.2. *If* $|\exp A_i - E| < 1$ $(i=1, 2, \cdots, m)$, *then* $\xi = \mu$.

To prove this, we need only observe that, by definition,

$$(\exp A_i)^{\alpha_{ik}/j} = \exp (\alpha_{ik} A_i/j) \quad \text{if} \quad |\exp A_i - E| < 1.$$

THEOREM 2.3. *The sets* $\bar{\lambda}$ *and* $\bar{\iota}$ *are identical.*

Since $\mathfrak{J} \subset \mathcal{L}$, $\iota \subset \lambda$. Given a matrix in $\mathcal{L}$ with coefficients $\rho_i(t)$, there exists a matrix in $\mathfrak{J}$, with coefficients $\mu_i(t)$, such that

$$\sum_{k=1}^{m} \int_a^b |\rho_k(t) - \mu_k(t)| \, |a_{ijk}| \, dt$$

is arbitrarily small $(i, j = 1, 2, \cdots, n)$. Therefore[8] $\iota \subset \bar{\lambda}$.

CARNEGIE INSTITUTE OF TECHNOLOGY

---

[8] See W. M. Whyburn, *Functional properties of the solutions of differential systems* Loc. cit.