EXPONENTIAL SOLUTIONS OF LINEAR SYSTEMS OF DIFFERENTIAL EQUATIONS WHOSE COEFFICIENT MATRIX IS SKEW SYMMETRIC

IRVING J. EPSTEIN

1. **Introduction and summary.** There exists a large literature on the exponential function of a matrix and on the representation of matrices as exponential functions of other matrices. For only one pair; A, E of matrices has the relation

$$E = \exp A$$

been thoroughly investigated, namely when A and E are both constant matrices.

Of much greater interest, however, are the cases where E or A represent a given set of matrices depending on one or several parameters, as for instance when E represents an arbitrary element of a Lie group. In these cases great difficulties arise if a "global" solution (for all values of the parameters) is sought for $E = \exp A$, where either E or A are given and the other one has to be found. Some of the difficulties encountered are, for instance, described in [1] and [2]. A basic reason for these difficulties is that the eigenvalues of a matrix $E(t)$ depending analytically on one parameter t are, in general, not analytic functions of t (as, for instance, in the case of

$$E(t) = \begin{pmatrix} t & t \\ 1 & -t \end{pmatrix}$$

at $t=0$). We present here a type of problem where a global solution always exists. Let $\Sigma(t)$ be a real skew symmetric $n \times n$ matrix depending analytically on the real variable t, and let the proper orthogonal matrix $O(t)$ be defined by the differential equation

$$\frac{d}{dt} O(t) = \Sigma(t) O(t), \qquad O(0) = I \tag{1.1}$$

where I is the unit matrix.

We shall show that the solution of equation (1.1), $O(t)$, can be expressed as

$$O(t) = e^{B(t)} \tag{1.2}$$

where $B(t)$ is an $n \times n$ real skew symmetric matrix whose elements are

Received by the editors May 13, 1965.

analytic functions of the real variable t for all t.

2. **We prove the following theorem.**

THEOREM. *Any proper real orthogonal matrix $O(t)$ whose elements are analytic functions of the real variable t for all t can be written as an exponential, $e^{B(t)}$, where $B(t)$ is a real skew symmetric matrix whose elements are also analytic in t for all t.*

Since every proper orthogonal matrix whose elements are analytic in t for all t satisfy equation (1.1) above, we see that the solutions of equation (1.1) which initially are proper orthogonal can be given by equation (1.2). For the proof of the theorem we shall need several lemmas.

LEMMA 1. *If $O(t)$ is a real $n\times n$ proper orthogonal matrix whose elements are analytic functions of the real variable t for all t, then the eigenvalues of $O(t)$ are also analytic in t for all t.*

PROOF. The eigenvalues all have absolute value one. Since our $O(t)$ is a real matrix, complex eigenvalues occur in pairs, λ and $\bar{\lambda}$ its conjugate. From this we see that our characteristic equation for $O(t)$ is a reciprocal equation. The substitution $\mu=\frac{1}{2}(\lambda+1/\lambda)$ reduces this equation to an equation of degree $n/2$ if n is even. If n is odd $\lambda(t)\equiv 1$ is always an eigenvalue. Dividing our equation by $(\lambda-1)$ there results again a reciprocal equation of even degree. In both cases we obtain an equation

$$\mu^l + c_1(t)\mu^{l-1} + \cdots + c_n(t) = 0 \tag{2.1}$$

whose coefficients $c_i(t)$ are analytic functions of t for all t and whose solutions are real and bounded for any real t. This follows from the fact that any solution of equation (2.1) is equal to $a(t)$ where $\lambda(t)=a(t)+ib(t)$, and $a(t)$ and $b(t)$ are real.

In the neighborhood of any fixed t_0 we have the following: the solutions of equation (2.1) are given by

$$\mu_i(t) = \mu_i(t_0) + \sum_{\nu=1}^{\infty} a_\nu^i (t-t_0)^{\nu/m_i}; \qquad i = 1, 2, \cdots, l. \tag{2.2}$$

Here $\mu_i(t_0)$ are the solutions of equation (2.1) corresponding to t_0. The a_ν^i are constant coefficients and m_i is a positive integer with $m_i \leq l$. Of course, m_i gives the multiplicity of the root $\mu_i(t_0)$. Let us write equation (2.2) as follows

$$\frac{\mu_i(t) - \mu_i(t_0)}{(t-t_0)^{1/m_i}} = a_1^i + \sum_{\nu=2}^{\infty} a_\nu^i (t-t_0)^{\nu-1/m_i}. \tag{2.3}$$

Now the left-hand side is real for real t. The summation on the right can be made arbitrarily small for $(t-t_0)$ sufficiently small. These two facts show us that a_1^i is real. In a similar manner we establish that all a_ν^i are real.

By $(t-t_0)^{1/m_i}$ we mean one determination of the m_ith root of $(t-t_0)$. All determinations are given by

$$\rho^j(t - t_0)^{1/m_i}, \qquad j = 0, 1, \cdots, m_i - 1$$

where ρ is a primitive m_ith root of unity. Since all a_ν^i are real this last requires that $a_\nu^i=0$ if ν is not a multiple of m_i. But then equation (2.2) tells us that the real part of $\lambda(t)$, $a(t)=\mu(t)$ is analytic in t for all t. Finally, since $b(t)=(1-a^2(t))^{1/2}$ and since $a(t)$ has a maximum at those values t where it is $+1$ and a minimum at those values of t where it is (-1) one can easily show that $b(t)$ is also analytic in t for all t. It follows that $\lambda(t)=a(t)+ib(t)$ is analytic in t for all t.

LEMMA 2. *To each eigenvalue $\lambda(t)$ we can determine an eigenvector $Z(t)=x(t)+iy(t)$ which is analytic in t for all t. Moreover, if $\lambda(t)$ is complex valued, then $x(t)$ and $y(t)$ are unit vectors which are orthogonal to each other for all t.*

PROOF. Let $\lambda(t)$ be a fixed eigenvalue. For each value of t it is a root of the characteristic equation, occurring with some multiplicity m. Let r denote the smallest value of the multiplicity m, occurring at some fixed t_0. Clearly $\lambda(t)$ has multiplicity r in some neighborhood of t_0. We now consider the following matrix:

$$(2.4) \qquad O - \lambda I = \begin{pmatrix} a_{11}(t) - \lambda(t) & a_{12}(t) & a_{1n}(t) \\ a_{21}(t) & a_{22}(t) - \lambda(t) & a_{2n}(t) \\ \cdots & \cdots & \cdots \\ a_{n1}(t) & a_{n2}(t) & a_{nn}(t) - \lambda(t) \end{pmatrix}.$$

If $r=n$ the result is trivial. We assume $r\neq n$. Since $\lambda(t)$ has multiplicity r in some neighborhood of t_0, we know that the rank of the matrix $O-\lambda I$ is $n-r$ in some neighborhood of t_0. (This follows from the fact that since O is orthogonal it is normal and for normal matrices the dimension of the null space of $(O-\lambda I)$ is equal to the multiplicity r of the eigenvalue $\lambda(t)$.) There is no loss of generality in assuming therefore that the determinant Δ formed from the first $(n-r)$ rows and columns of the matrix $(O-\lambda I)$ does not vanish in some neighborhood of t_0.

We consider now the following system of equations

$$
(2.5)\qquad
\begin{aligned}
&(a_{11}-\lambda)W_1 + a_{12}W_2 + \cdots + a_{1,n-r}W_{n-r} + \cdots + a_{1n}W_n = 0\\
&a_{21}W_1 + (a_{22}-\lambda)W_2 + \cdots + a_{2,n-r}W_{n-r} + \cdots + a_{2n}W_n = 0\\
&\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\\
&a_{n-r,1}W_1 + a_{n-r,2}W_2 + \cdots + (a_{n-r,n-r}-\lambda)W_{n-r} + \cdots\\
&\qquad\qquad\qquad\qquad\qquad\qquad\qquad + a_{n-r,n}W_n = 0.
\end{aligned}
$$

Since $\Delta(t)\neq 0$ in some neighborhood of t_0, we can solve for W_1, W_2, $\cdots$, W_{n-r} in terms of the remaining W's. If we now set $W_{n-r+1} = W_{n-r+2} = \cdots = W_n = \Delta$ then we have a vector $W = (W_1, W_2, \cdots, W_n)$ which does not vanish in some neighborhood of t_0 and is analytic everywhere. Since each of the last r rows of the matrix $(O-\lambda I)$ is a linear combination of the first $n-r$ rows we see that the vector $W = (W_1, \cdots, W_n)$ is an eigenvector of $(O-\lambda I)$ in some neighborhood of t_0. Now $(O-\lambda I)$ is analytic everywhere. Also the vector W is analytic everywhere. Hence the relation $(O-\lambda I)W = 0$ which holds in some neighborhood of t_0 holds in fact everywhere. The vector W cannot vanish on any interval as this would imply that it vanishes identically. However, W can vanish at isolated points. At these values of t where W vanishes it could not be regarded as an eigenvector.

We consider now the vector $\tilde{W}$ given by

$$\tilde{W} = \left(\frac{W_1}{\rho}, \frac{W_2}{\rho}, \cdots, \frac{W_n}{\rho}\right)$$

where

$$\rho = \left\{\sum_{i=1}^{n} (u_i^2 + v_i^2)\right\}^{1/2}$$

and where u_i and v_i are the real and imaginary parts respectively of W_i.

We show that the vector $\tilde{W}$ is never zero. It is clear that $\tilde{W}$ is not zero at those values of t where W is not zero. We consider $\tilde{W}$ at some value t where W vanishes. There is no loss of generality in assuming this value of t to be $t=0$.

Since each u_i and v_i are analytic everywhere, we have

$$u_i(t) = \sum_{j=l_i}^{\infty} a_j^{(i)} t^j \text{ where } a_{l_i}^{(i)} \neq 0; \qquad i = 1, 2, \cdots, n,$$

$$v_i(t) = \sum_{j=m_i}^{\infty} b_j^{(i)} t^j \text{ where } b_{m_i}^{(i)} \neq 0; \qquad i = 1, 2, \cdots, n.$$

Let

$$\gamma = \min_{i} \left\{ a_{l_i}^{(i)}, b_{m_i}^{(i)} \right\}.$$

There is no loss of generality in assuming that $a_{l_1}^{(1)}$ is our minimum.

We consider the first component of the vector $\tilde{W}$, namely W_1/ρ. We have

$$\frac{W_1}{\rho} = \frac{u_1 + iv_1}{t^{l_1}((a_{l_1}^{(1)})^2 + \cdots)^{1/2}}.$$

Clearly the real part of this last expression is not zero since $a_{l_1}^{(1)} \neq 0$ and the square root is different from zero. This proves that the vector $\tilde{W}$ is never zero. Like W we see that $\tilde{W}$ is analytic everywhere and moreover it is an eigenvector of $O - \lambda I$ for all t.

If $\lambda(t)$ is complex valued for $t = t_1$ then it follows[1] that the real and imaginary parts $\tilde{u} = (\tilde{u}_1, \cdots, \tilde{u}_n)$ and $\tilde{v} = (\tilde{v}_1, \cdots, \tilde{v}_n)$ of $\tilde{W}$ are orthogonal to one another and of equal length $(\{\sum \tilde{u}_i^2\}^{1/2} = \{\sum \tilde{v}_i^2\}^{1/2})$ in some neighborhood of t_1. Since $\tilde{u}$ and $\tilde{v}$ are analytic everywhere these relationships hold for all t. It is clear that in this case the vector $Z(t) = x(t) + iy(t)$ with

$$x(t) = \frac{\tilde{u}(t)}{l} \quad \text{and} \quad y(t) = \frac{\tilde{v}(t)}{l} \quad \text{where} \quad l = \left(\sum \tilde{u}_i^2\right)^{1/2} = \left(\sum \tilde{v}_i^2\right)^{1/2}$$

meets the requirements of Lemma 2.

If $\lambda(t)$ is always real then the same argument establishes the existence for all t of a real unit eigenvector $x(t)$.

LEMMA 3. *Let $\alpha_1(t), \alpha_2(t), \cdots, \alpha_s(t)$ be $s < n$ real orthonormal vectors which are analytic in t for all real t. There exist $n-s$ additional real vectors $\alpha_{s+1}, \cdots, \alpha_n$ such that the set of n vectors $\alpha_1, \cdots, \alpha_n$ are an orthonormal set which are analytic in t for all real t.*

PROOF. We show how to obtain the vector α_{s+1}. Let $\beta(t) = (\beta_1, \cdots, \beta_n)$. Consider the matrix given by

$$\begin{pmatrix} \alpha_{11} & \alpha_{12} & \cdots & \alpha_{1n} \\ \alpha_{21} & \alpha_{22} & \cdots & \alpha_{2n} \\ \cdot & \cdot & \cdot & \cdot \\ \alpha_{s1} & \alpha_{s2} & \cdots & \alpha_{sn} \end{pmatrix} \tag{2.6}$$

where $\alpha_i(t) = (\alpha_{i1}, \alpha_{i2}, \cdots, \alpha_{in})$; $i = 1, 2, \cdots, s$.

[1] S. Perlis, *Theory of matrices*, Addison-Wesley, Cambridge, Mass., 1952, p. 200.

The rank of this matrix is s since the s vectors $\alpha_1, \cdots, \alpha_s$ are linearly independent for all t. We assume then that the determinant d formed from the first s rows and columns of this matrix is not zero in some neighborhood, say t_0, of t. The system of equations

$$\begin{aligned} \alpha_{11}\beta_1 + \alpha_{12}\beta_2 + \cdots + \alpha_{1n}\beta_n &= 0 \\ \cdots\cdots\cdots\cdots & \\ \alpha_{s1}\beta_1 + \alpha_{s2}\beta_2 + \cdots + \alpha_{sn}\beta_n &= 0 \end{aligned} \tag{2.7}$$

can then be solved for $\beta_1, \beta_2, \cdots, \beta_s$ in terms of the remaining β's in some neighborhood of t_0. Set $\beta_{s+1} = \beta_{s+2} = \cdots = \beta_n = d$. We obtain $\beta = (\beta_1, \cdots, \beta_n)$ which is analytic everywhere and does not vanish in some neighborhood of t_0. The vector β/l where $l = (\sum \beta_i^2)^{1/2}$ is never zero, is analytic in t for all t, is of unit length, and is orthogonal to $\alpha_1, \cdots, \alpha_s$ for all t. This proves Lemma 3.

With the aid of these lemmas the proof of our theorem can now be established by induction. It is easy to show that any 2×2 real proper orthogonal matrix $O_2(t)$ whose elements are analytic in t for all t is orthogonally similar to

$$\tilde{O}_2(t) = \begin{pmatrix} a(t) & b(t) \\ -b(t) & a(t) \end{pmatrix}.$$

We have $P^{-1}(t)O_2(t)P(t) = \tilde{O}_2(t)$ where $P(t)$ is real and analytic in t and proper orthogonal and where a and b are real and analytic for all t and of course $a^2+b^2=1$. Assume that any real proper orthogonal matrix of order τ less than n whose elements are analytic for all t is orthogonally similar to $\tilde{O}_\tau(t) = \operatorname{diag}(R_q, C_1, \cdots, C_m)$ where $R_q = \operatorname{diag}(r_1, \cdots, r_q)$ is real and

$$C_i = \begin{pmatrix} a_i & b_i \\ -b_i & a_i \end{pmatrix}; \qquad i = 1, 2, \cdots, m.$$

We assume therefore that there is a real proper orthogonal matrix $P(t)$ whose elements are analytic for all t such that $P^{-1}O_\tau P = \tilde{O}_\tau$. The elements r_j, a_i, b_i are real and analytic for all t. We have $\tau = q + 2m$.

Now consider the matrix $O_n(t)$ where n is its order. By Lemma 2 there exists, if $\lambda(t) = a_1 + ib_1$ is complex valued for some t, an $x(t)$ and $y(t)$ which are real and analytic and orthonormal for all t. By Lemma 3 we can find $n-2$ additional vectors which are analytic in t and which, together with x and y, form a set of n real analytic orthonormal vectors.

Construct the matrix S with these vectors as the column vectors and so chosen that det $S = 1$. We have

$$S^{-1}O_nS = \begin{pmatrix} A & 0 \\ 0 & O_{n-2} \end{pmatrix} \quad \text{where} \quad A = \begin{pmatrix} a_1 & b_1 \\ -b_1 & a_1 \end{pmatrix}$$

and where O_{n-2} is real proper orthogonal and analytic for all t. By the induction assumption there exists a Q_{n-2} which is real proper orthogonal and analytic for all t such that $Q_{n-2}^{-1}O_{n-2}Q_{n-2}=\tilde{O}_{n-2}$. The matrix

$$Q_n = \begin{pmatrix} I_2 & 0 \\ 0 & Q_{n-2} \end{pmatrix} \quad \text{with} \quad I_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

is proper orthogonal and analytic for all t. Set $P_n=SQ_n$ we have $P_n^{-1}O_nP_n = Q_n^{-1}S^{-1}O_nSQ_n = \tilde{O}_n = \operatorname{diag}(R_s, C_1, \cdots, C_t)$ where $R_s = \operatorname{diag}(r_1, r_2, \cdots, r_s)$ and

$$C_i = \begin{pmatrix} a_i & b_i \\ -b_i & a_i \end{pmatrix} \qquad \text{for } i = 1, 2, \cdots, t.$$

Of course $s+2t=n$ and a_i and b_i are the real and imaginary parts of the complex eigenvalue $\lambda_i(t)$. Now each

$$C_i = e^{\beta_i J} \quad \text{where} \quad J = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$$

and $\beta(t)$ is analytic in t for all t. If (-1) is an eigenvalue of a proper orthogonal matrix then it has even multiplicity. Hence

$$R_s = \operatorname{diag}(e^{\pi J}, e^{\pi J}, \cdots, e^{\pi J}, e^0, e^0, \cdots, e^0).$$

(If -1 is not an eigenvalue we omit $e^{\pi J}$ and if $+1$ is not an eigenvalue we omit e^0.)

If follows that

$$\begin{aligned} \tilde{O}_n &= \operatorname{diag}(e^{\pi J}, \cdots, e^{\pi J}, e^0, \cdots, e^0, e^{\beta_1 J}, \cdots, e^{\beta_t J}) \\ &= \exp[\operatorname{diag}(\pi J, \cdots, \pi J, 0, \cdots, 0, \beta_1 J, \cdots, \beta_t J)] = e^T \end{aligned}$$

where $T=\operatorname{diag}(\pi J, \cdots, \pi J, 0, \cdots, 0, \beta_1 J, \cdots, \beta_t J)$.

Since $P_n^{-1}O_nP_n=\tilde{O}_n=e^T$, we get $O_n=e^{P_nTP_n^{-1}}$. With $\beta=P_nTP_n^{-1}$ this establishes our theorem.

References

1. G. Bachman and M. J. Hillman, *On the parametrization of the proper orthogonal groups*, Arch. Math. 10 (1959), 93–100.

2. W. Magnus, *A Fourier theorem for matrices*, Proc. Amer. Math. Soc. 6 (1955), 880–891.

United States Army Signal Research and Development Laboratory,
Fort Monmouth, New Jersey