HAUSDORFF-ANALYTIC FUNCTIONS OF MATRICES\(^1\)

WALTER O. PORTMANN\(^2\)

1. Introduction. The concept of Hausdorff-analytic (\(H\)-analytic) functions on an (associative) hypercomplex system, over the complex number field, with an identity \([1]\) is, in this paper, applied to functions of matrices with distinct eigenvalues. In the process of showing a sufficient condition for a matrix function to be \(H\)-analytic in a neighborhood of a matrix \(Z_0\) with distinct eigenvalues, it is shown that the projections (Frobenius covariants) are \(H\)-analytic in a neighborhood of \(Z_0\), and that there exists a matrix function \(Q = Q(Z)\) which is \(H\)-analytic in a neighborhood of \(Z_0\) such that \(Q^{-1}ZQ = \Lambda = \text{diag} (\lambda_k)\), \(\lambda_i \neq \lambda_j\) for \(i \neq j\), for all \(Z\) in that neighborhood of \(Z_0\).

It has previously been shown \([2]\) that if \(f(Z)\) is a function on \(\mathfrak{M}\) (the algebra of all square matrices of order \(n\) over the complex number field) whose component functions are analytic functions, in some open domain, of the complex (component) variables \(z_{ij}\) of \(Z = (z_{ij})\), then \(f(Z)\) is \(H\)-analytic in a corresponding open domain of \(\mathfrak{M}\).

2. \(H\)-analyticity of the projections of a matrix with distinct eigenvalues.

**Lemma 2.1.** The projections (or Frobenius covariants) \(P_k(Z)\) corresponding to the eigenvalues \(\lambda_k\), \(k = 1, \cdots, s\), of a matrix \(Z\) in \(\mathfrak{M}\) are given by

\[
P_k(Z) = \frac{1}{2\pi i} \int_{C_k} (\lambda I - Z)^{-1} d\lambda,
\]

where \(C_k\) is a circle in the complex \(\lambda\)-plane containing \(\lambda_k\) but none of the other \(\lambda_j\).

**Proof.** By \([4, p. 22]\),

\[
(\lambda I - Z)^{-1} = \sum_{j=1}^{s} P_j(Z) \sum_{m=0}^{r_j-1} \frac{(Z - \lambda_j I)^m}{(\lambda - \lambda_j)^{m+1}},
\]

where \(r_j\) is the index of \(\lambda_j\).

The integral, over a curve in the complex \(\lambda\)-plane, of a matrix of complex functions is given by the matrix of integrals, over the curve, of the matrix elements. Therefore if we let \(C_k\) be a circle containing

---

\(^1\) This paper was prepared under the facilities granted by the Case Research Fund.

\(^2\) The author is now associated with Arizona State University.
\( \lambda_k \) but none of the other \( \lambda_j \), then
\[
\frac{1}{2\pi i} \int_{c_k} (\lambda I - Z)^{-1} d\lambda = \frac{1}{2\pi i} \sum_{j=1}^{s} P_j(Z) \sum_{m=0}^{r_j-1} (Z - \lambda_j I)^m \int_{c_k} \frac{d\lambda}{(\lambda - \lambda_j)^{m+1}} = P_k(Z)
\]

We will use this representation to prove the following.

**Theorem 2.1.** If \( Z_0 \) is a matrix with distinct eigenvalues, then there exists a neighborhood \( N \) of \( Z_0 \) such that for \( Z \) in \( N \), \( P_k(Z) \) is an \( H \)-analytic function of \( Z \).

**Proof.** Let \( \lambda_k^0, \; k = 1, \ldots, n \), be the (distinct) eigenvalues of \( Z_0 \), then
\[
P_k(Z_0) = \frac{1}{2\pi i} \int_{c_k} (\lambda I - Z)^{-1} d\lambda
\]
is the projection of \( Z_0 \) corresponding to \( \lambda_k^0 \), where \( C_k \) is a sufficiently small circle which has \( \lambda_k^0 \) as its center and all other \( \lambda_0^0 \) in its exterior, that is, if \( |\lambda_k^0 - \lambda_j^0| > 2\epsilon \) for \( j \neq k \), then \( C_k \): \( |\lambda - \lambda_k^0| = \epsilon \) will be sufficient.

Now, for all matrices \( Z \) sufficiently near \( Z_0 \), that is, such that \( \text{norm}(Z - Z_0) \) is sufficiently small (where, for convenience, the norm of any matrix \( X = (x_{ij}), \; i = 1, \ldots, m, \; j = 1, \ldots, n \), with complex components, shall be defined by \( \text{norm}(X) = \max_{i,j} |x_{ij}| \), the eigenvalues \( \lambda_1, \ldots, \lambda_n \) of \( Z \) will be near those of \( Z_0 \), since the eigenvalues of a matrix are continuous functions of the elements of the matrix \([2]\). Thus, a neighborhood \( N \) of \( Z_0 \) may be chosen such that \( |\lambda_k - \lambda_k^0| < \epsilon, \; k = 1, \ldots, n \), then for each \( Z \) in \( N \), \( Z \) has distinct eigenvalues \( \lambda_j \), and \( \lambda_k \) lies within \( C_k \) while all other \( \lambda_j \) lie outside \( C_k \). Hence for all \( Z \) in \( N \),
\[
P_k(Z) = \frac{1}{2\pi i} \int_{c_k} (\lambda I - Z)^{-1} d\lambda.
\]
The \( r, s \) element of the matrix \( P_k(Z) \) is given by
\[
P_k(Z)_{rs} = \frac{1}{2\pi i} \int_{c_k} R_{rs}(\lambda, z_{ij}) d\lambda,
\]
where \( R_{rs}(\lambda, z_{ij}) \) is the quotient of two polynomials in \( \lambda \) and the \( z_{ij} \), \( i, \; j = 1, \ldots, n \). Since \( C_k \) does not pass through any of the zeros of \( \text{det}(\lambda I - Z) \), regardless of what \( Z \) in \( N \) is chosen, \( R_{rs}(\lambda, z_{ij}) \) is a continuous function of the complex variables \( \lambda \) and \( z_{ij}, i, \; j = 1, \ldots, n \).
where each \( z_{ij} \) ranges over a region \( N_{ij} \) determined by \( N \), and \( \lambda \) lies on \( C_k \); also \( R_{ij}(\lambda, z_{ij}) \) is an analytic function of each \( z_{ij} \) in \( N_{ij} \), for every value of \( \lambda \) of \( C_k \). Therefore, \( P_k(Z)_{rs} \) is an analytic function of each \( z_{ij} \) of \( Z \) in \( N \), and hence, the components \( P_k(Z)_{rs} \) are analytic functions of \( z_{ij} \), that is \( P_k(Z) \) is \( H \)-analytic at \( Z \) in \( N \).

3. The existence of \( Q(Z) \).

**Lemma 3.1.** If \( Z_0 \) is a matrix with distinct eigenvalues and \( x \) is a vector such that \( P_k(Z_0)x \neq 0 \) for every \( k = 1, \ldots, n \), then for \( Z \) sufficiently near \( Z_0 \), \( P_k(Z)x \neq 0 \), \( k = 1, \ldots, n \).

**Proof.** Since \( P_k(Z_0)x \neq 0 \) for every \( k = 1, \ldots, n \), there exists a \( \delta > 0 \) such that \( \text{norm}(P_k(Z_0)x) \geq \delta \) for every \( k \). Now \( P_k(Z) \) is an \( H \)-analytic function of \( Z \) in a neighborhood of \( Z_0 \) and therefore a continuous function of \( Z \) in a neighborhood of \( Z_0 \). Thus, for \( Z \) near \( Z_0 \), \( P_k(Z) \) is near \( P_k(Z_0) \) and therefore \( P_k(Z)x \) is near \( P_k(Z_0)x \); in particular \( Z \) may be chosen sufficiently close to \( Z_0 \) such that

\[
\text{norm}(P_k(Z_0)x - P_k(Z)x) < \delta/2,
\]

for all \( k \). Hence, for all \( Z \) in such a neighborhood \( N \) of \( Z_0 \),

\[
\text{norm}(P_k(Z)x) \geq \text{norm}(P_k(Z_0)x) - \text{norm}(P_k(Z_0)x - P_k(Z)x) > \delta - \delta/2 = \delta/2 > 0.
\]

Thus, for \( Z \) in \( N \), \( P_k(Z)x \neq 0 \) for every \( k = 1, \ldots, n \).

**Theorem 3.1.** Let \( Z_0 \) be a matrix with distinct eigenvalues; then there exists a nonsingular matrix \( Q \) whose components are analytic functions of the elements of \( Z \), for \( Z \) in some neighborhood \( N \) of \( Z_0 \) (therefore \( Q \) and \( Q^{-1} \) are \( H \)-analytic functions of \( Z \) in \( N \)), such that \( Q^{-1}ZQ = \Lambda = \text{diag}(\lambda_k) \) for all \( Z \) in \( N \).

**Proof.** Choose a vector \( x \) such that \( P_k(Z_0)x \neq 0 \) for all \( k = 1, \ldots, n \), then by Lemma 3.1, for \( Z \) sufficiently near \( Z_0 \), \( Q_k = P_k(Z)x \neq 0 \). By [4, p. 22], \((Z - \lambda_k I)Q_k = 0\), since the index \( r_k \) of \( \lambda_k \) is 1 for all \( k \), that is, \( ZQ_k = \lambda_k Q_k \). Now, let \( Q \) be the matrix whose \( j \)th column is \( Q_j \), then \( ZQ = QA \), where \( A = \text{diag}(\lambda_k) \). Therefore, since the \( Q_k \) are linearly independent \( Q^{-1}ZQ = \Lambda \).

Since \( Q_k = P_k(Z)x \), the components of \( Q_k \) are linear combinations of the elements of \( P_k(Z) \) and therefore, by Theorem 2.1, they are analytic functions of the elements \( z_{ij} \) of \( Z \) in a neighborhood of \( Z_0 \). Hence \( Q \) is \( H \)-analytic in a sufficiently small neighborhood \( N \) of \( Z_0 \).

Also \( Q^{-1} = (S_{ij}/\text{det}(Q)) \), where \( S_{ij} \) is the cofactor of the \( j, i \) component of \( Q \). Since \( \text{det}(Q) \neq 0 \), and \( S_{ij} \) and \( \text{det}(Q) \) are polynomials
with constant coefficients in the analytic components of \( Q \), the components at \( Q^{-1} \) are analytic functions of the elements of \( Z \) in \( N \) and therefore \( Q^{-1} \) is also \( H \)-analytic in \( N \).

Note: If one knows a \( Q_0 \) such that \( Q_0^{-1}Z_0Q_0 = \Lambda_0 = \text{diag}(\lambda_i^0) \) for the given \( Z_0 \), then, for the above \( x \), one may choose \( x = \sum_{j=1}^n Q_j^{(p)} \), where \( Q_j^{(p)} \) is the \( j \)th column of \( Q_0 \), since \( Z_0Q_0^{(p)} = \lambda_i^0Q_0^{(p)} \); that is, \( P_k(Z_0)Q_0^{(p)} = \delta_{jk}Q_0^{(p)} \), and therefore \( P_k(Z_0)x = Q_0^{(p)} \neq 0 \) for every \( k = 1, \ldots, n \).

**Corollary 3.1.** If the matrix \( Z \) has distinct eigenvalues \( \lambda_k \), \( k = 1, \ldots, n \), then the \( \lambda_k \) are analytic functions of the components \( z_{ij} \) of \( Z \).

### 4. A sufficient condition for \( H \)-analyticity of a matrix function at a matrix with distinct eigenvalues

The hypothesis of (i) of the following theorem would be desirable if one wished to view this as a theory for functions of linear transformations of a finite dimensional vector space, for the invariance under similarity transformations, \( F(Y) = P^{-1}F(X)P \) for \( Y = P^{-1}XP \), permits the definition of a function of a finite dimensional linear transformation to be independent of the choice of basis for the vector space.

**Theorem 4.1.** Let \( F = \sum_{i,j=1}^n f_{ij}E_{ij} \) be a matrix function (where \( E_{ij} \) is the \( n \times n \) matrix with a 1 in the \( i, j \) position and zeros elsewhere).

(i) Let \( F \) be such that \( F \) defined at \( X \) and \( Y = P^{-1}XP \) implies that \( F \) is defined at \( Y \) and \( F(Y) = P^{-1}F(X)P \); then, for a given \( X \) at which \( F \) is defined, \( F(X) \) is a polynomial in \( X \). In particular, if \( \Lambda = \text{diag}(\lambda_i) \) is a diagonal matrix at which \( F \) is defined, then \( F(\Lambda) \) is a diagonal matrix, that is,

\[
\left[ f_{ij} \right]_{\sum_{r=1}^n \delta_{ir} \lambda_r} = \delta_{ij}g_i(\lambda_1, \ldots, \lambda_n).
\]

(ii) Further, if \( F \) is also such that the diagonal component functions \( f_{ii} \) are analytic functions of each \( \lambda_j \) at the components of a diagonal matrix with distant eigenvalues \( \lambda_i \), at which \( F \) is defined, that is,

\[
\left. \frac{\partial f_{ij}}{\partial z_{jj}} \right|_{\sum_{r=1}^n \delta_{ir} \lambda_r} = \frac{\partial g_i}{\partial \lambda_j}
\]

exist for \( i, j = 1, \ldots, n \), then, \( F \) is \( H \)-analytic at a matrix \( Z \) with distinct eigenvalues at which \( F \) is defined.

**Proof.** To prove the first part of this theorem we shall use the following lemma obtained by Richter [3].

**Lemma 4.1.** Let \( F \) be a function which satisfies the hypothesis of (i) of Theorem 4.1, and let \( X \) be a matrix for which \( F(X) \) is defined. If
$XB = BX$, then $F(X)B = BF(X)$, that is, $F(X)$ commutes with every matrix which commutes with $X$.

Any matrix $F(X)$ satisfying the conclusion of the above lemma is a polynomial in $X$ [5].

To prove the second part of the theorem, let $F$ be defined at a matrix $Z$ with distinct eigenvalues. By Theorem 3.1 there exists a $Q = Q(Z)$, $H$-analytic at $Z$, such that $Q^{-1}ZQ = \Lambda = \text{diag}(\lambda_i)$, where the $\lambda_i$ are the eigenvalues of $Z$; therefore by (i), $F(Z) = QF(\Lambda)Q^{-1}$ and

$$F(\Lambda) = \sum_{i=1}^{n} g_i E_{ii}. $$

Therefore, since the $g_i$ are analytic functions of the $\lambda_i$ and by Corollary 3.1, the $\lambda_i$ are analytic functions of the $z_{rs}$, $r, s = 1, \cdots, n$, the $f_{ij}$ are analytic functions (in a neighborhood) of the components $z_{rs}$ of $Z$. Thus $F$ if $H$-analytic at $Z$.

REFERENCES

5. H. W. Turnbull and A. C. Aitken, An introduction to the theory of canonical matrices, Blackie and Son Ltd., 1932, Chap. X.

Case Institute of Technology