

Über Matrixfunktionen.

Von

H. RICHTER in Haltingen (Baden).

§ 1. Einleitung.

Es kommt in den physikalischen Anwendungen oft vor, daß die reell- oder komplexwertigen Komponenten einer quadratischen n -reihigen Matrix B eindeutige stetige Funktionen sind der Komponenten einer anderen quadratischen Matrix A des gleichen Grades, derartig, daß diese Abhängigkeit invariant ist gegen affine Änderung des Koordinatensystems. Ein Beispiel dieser Art wird durch die Abhängigkeit des Spannungstensors vom Deformationstensor in der allgemeinen Elastizitätstheorie isotroper Medien gegeben, wenn diese beiden Tensoren geeignet gewählt werden¹⁾. Andererseits weiß man, daß es möglich ist, ganze analytische Funktionen auf Matrizen anzuwenden, da man durch formales Einsetzen einer Matrix A anstelle der unabhängigen Variablen stets eine konvergente Reihe erhält, und daß diese Art der Abhängigkeit ebenfalls gegen die Transformation mit einer beliebigen dritten Matrix invariant ist, so daß diese funktionelle Abhängigkeit einen Spezialfall der zuerst genannten Abhängigkeit bildet. Es erscheint daher zweckmäßig, zunächst allgemein nach der Gestalt der invarianten Abhängigkeit zweier Matrizen zu fragen und erst dann zu untersuchen, wann diese Abhängigkeit als Anwendung einer gewöhnlichen Funktion auf Matrizen aufgefaßt werden kann. Dementsprechend unterscheiden wir im folgenden die Begriffe der „allgemeinen Matrixfunktion“ und der „gewöhnlichen Matrixfunktion“. Es wird sich dabei zeigen, daß diese Unterscheidung nur dann wesentlich ist, wenn wir eine bestimmte funktionelle Abhängigkeit auf variabel gedachtes A anwenden; nicht dagegen, wenn wir die Gesamtheit aller möglichen Funktionen eines festen A betrachten. Von den gewöhnlichen Matrixfunktionen dürften das größte Interesse diejenigen beanspruchen, die durch analytische (jedoch nicht notwendig ganze) Funktionen vermittelt werden: „analytische Matrixfunktionen“. Wir werden in § 2 zunächst den Begriff der allgemeinen eindeutigen Matrixfunktion klären und Normaldarstellungen dafür angeben. Für spezielle A wird die Matrixfunktion durch einen Grenzprozeß definiert, was beim Zusammenfallen von Eigenwerten von A zu Mehrdeutigkeiten Anlaß geben kann.

In § 3 wird die Anwendung gewöhnlicher und insbesondere analytischer Funktionen $f(x)$ auf Matrizen besprochen. Es zeigt sich, daß bei analytischem $f(x)$ der Wert von $f(A)$ auch für die A mit mehrfachen Eigenwerten eindeutig bestimmt ist und bei Kenntnis der Eigenwerte von A und ihrer Vielfachheiten in einfacher Weise berechnet werden kann. Eine notwendige und hinreichende Bedingung für die Realität von $f(A)$ bei reellem A wird abgeleitet.

¹⁾ RICHTER, H.: Verzerrungstensor, Verzerrungsdeviator und Spannungstensor bei endlichen Formänderungen. Z. ang. Math. Mech. **29**, 65–75 (1949).

§ 4 behandelt die Lösung der Matrixgleichung $A = \varphi(B)$ bei bekanntem A und eindeutigem analytischen φ . Die allgemeine Gestalt von B wird angegeben und für den Fall $\varphi(x) = x^2$, also $B = \sqrt{A}$ demonstriert.

In § 5 wird schließlich für analytische Matrixfunktionen $f(A)$ der Zusammenhang des Differentials mit $f'(A)$ behandelt und ein Analogon zu der gewöhnlichen Beziehung $df = f' \cdot dx + o((dx)^2)$ abgeleitet.

Bezeichnungen. Zur Matrix A mit den Komponenten $a_{ik} = (A)_{ik}$ ist $\bar{A}$ die an der Hauptdiagonale Gespiegelte, $|A|$ die Determinante, $\{A\}$ die Spur. Für die Eigenwerte λ_v von A sind die Potenzsummen: $s_m = \sum_v \lambda_v^m = \{A^m\}$.

E ist die Einheitsmatrix. (e_v) ist eine Diagonalmatrix mit den Elementen e_v . n ist der Grad von A .

In Übereinstimmung mit der üblichen Konvention beim Einsetzen von Matrizen in Polynome setzen wir $A^0 = E$ für jedes A .

a und a^* sind konjugiert-komplexe Zahlen.

§ 2. Die allgemeine stetige Matrixfunktion.

Sind vermöge einer Rechenvorschrift die Komponenten der Matrix B eindeutige stetige Funktionen der variabel gedachten Komponenten der Matrix A in einem offenen Teilbereich mit teilweise hinzugenommenen Randpunkten des n^2 -dimensionalen Raumes der Komponenten von A , dann bedeutet es keine wesentliche Einschränkung, die Randpunkte zunächst wegzulassen, da die Werte von B auf denselben stetig aus inneren Stellen gewonnen werden können. Wir können aus dem gleichen Grunde die geringerdimensionale Menge aller A streichen, die zusammenfallende Eigenwerte haben. Wir bilden daher die folgende

Definition 1: Es sei $\mathfrak{B}_0$ ein offener Bereich von Matrizen A , die lauter verschiedene Eigenwerte haben. Mit A seien auch alle äquivalenten CAC^{-1} zu A in $\mathfrak{B}_0$ enthalten. Zu jedem $A \in \mathfrak{B}_0$ sei stetig ein B definiert. Dann heißt B eine allgemeine stetige Matrixfunktion von A in $\mathfrak{B}_0$, symbolisch

$$(2.1) \quad B = Fu(A),$$

wenn für jedes C mit $|C| \neq 0$ gilt:

$$(2.2) \quad CBC^{-1} = Fu(CAC^{-1}).$$

Man sieht, daß der Begriff der allgemeinen Matrixfunktion eine Verallgemeinerung des Begriffes der Invariante darstellt.

Um nun die oben ausgeschlossenen A mitzuerfassen, ergänzen wir durch

Definition 2: Liegt A in der abgeschlossenen Hülle von $\mathfrak{B}_0$, dann gehört A mit dem Funktionswert $Fu(A) = B$ zum Definitionsbereich $\mathfrak{B}$ der Matrixfunktion, wenn es eine stetige Folge²⁾ $A(t)$, $t > 0$, gibt mit:

$$A(t) \in \mathfrak{B}_0; \quad \lim_{t \rightarrow 0} A(t) = A; \quad \lim_{t \rightarrow 0} Fu(A(t)) = B.$$

Hierbei kann es vorkommen, daß $Fu(A)$ abhängig wird von der definierenden Folge $A(t)$. Wir unterscheiden daher gemäß

²⁾ Hierbei kann an eine stetige Abhängigkeit der A von dem reellen Parameter t oder auch an eine konvergente Folge $A(t_n)$ mit $t_n \rightarrow 0$ gedacht werden.

Definition 3: Ein $A \in \mathfrak{B}$ ist eine *reguläre Stelle* der Matrixfunktion, wenn in Definition 2 B unabhängig ist von der Folge $A(t)$.

Anderenfalls ist A eine *richtungssinguläre Stelle*.

Da $Fu(A)$ in $\mathfrak{B}_0$ stetig erklärt war, ist automatisch jedes $A \in \mathfrak{B}_0$ eine reguläre Stelle.

Beispiel 1: Für $Fu(A) = \frac{A \cdot |A|}{|A|}$ besteht $\mathfrak{B}_0$ aus allen Matrizen mit lauter verschiedenen Eigenwerten und $|A| \neq 0$. $\mathfrak{B}$ enthält alle Matrizen. Jedes A ist eine reguläre Stelle mit $Fu(A) = A$.

Beispiel 2: $Fu(A) = \sqrt[n]{|A|^{-1}} \cdot A$ hat das gleiche $\mathfrak{B}_0$ wie in Beispiel 1. Ist ein A beliebig mit $|A| \neq 0$, dann erhalten wir aus Definition 2 unabhängig von $A(t)$ stets $B = \sqrt[n]{|A|^{-1}} \cdot A$. Ist $A \neq 0$ mit $|A| = 0$, so existiert $\lim_{t \rightarrow 0} Fu(A(t))$ für keine Folge $A(t)$. Ist schließlich $A = 0$, so kann aus

Definition 2 für B eine beliebige Matrix mit der Determinante 1 erhalten werden. $\mathfrak{B}$ besteht dann aus allen Matrizen mit $|A| \neq 0$ und der Matrix $A = 0$, wobei die letztere die einzige richtungssinguläre Stelle ist.

Wir können nun Definition 2 abrunden durch

Satz 1: Sei in $\mathfrak{B}$ enthalten eine stetige Folge $A(t)$ mit je einem $Fu(A(t))$, $\lim_{t \rightarrow 0} A(t) = A$ und $\lim_{t \rightarrow 0} Fu(A(t)) = B$. Dann ist $A \in \mathfrak{B}$ mit $Fu(A) = B$.

Beweis: Nach Definition 2 gibt es $A(t, u) \in \mathfrak{B}_0$ mit $\lim_{u \rightarrow 0} A(t, u) = A(t)$ und $\lim_{u \rightarrow 0} Fu(A(t, u)) = Fu(A(t))$, $Fu(A(t, u))$ gemäß Definition 1. Wir wählen zu jedem t ein $u(t) > 0$ mit $u(t) \leq t$ und $\max_{i,k} |(A(t, u) - A(t))_{ik}| \leq t$ nebst $\max_{i,k} |(Fu(A(t, u)) - Fu(A(t)))_{ik}| \leq t$ für $u \leq u(t)$. $A(t) = A(t, u(t))$ erfüllt dann Definition 2.

Ebenso unmittelbar folgt aus unseren Definitionen, daß bei $A \in \mathfrak{B}$ mit $B = Fu(A)$ auch jede Transformierte $A_1 = CAC^{-1}$ in $\mathfrak{B}$ liegt mit $Fu(A_1) = CBC^{-1}$. Dabei sind die Transformaten zu regulären (richtungssingulären) Stellen wieder reguläre (richtungssinguläre) Stellen in $\mathfrak{B}$. Hieraus ergibt sich weiter, daß wir für die Untersuchung von $Fu(A)$ stets A in der Normalform

$$(2.3) \quad A' = \begin{pmatrix} A_1 & & 0 \\ & A_2 & \\ & & \ddots \\ 0 & & & A_m \end{pmatrix} \quad \text{mit } A_k = \begin{pmatrix} \lambda_k & 1 & & 0 \\ & \lambda_k & 1 & \\ & & \ddots & \\ 0 & & & \lambda_k \end{pmatrix}$$

annehmen können, in die bekanntlich jedes A affin transformiert werden kann.

Ist speziell C mit A vertauschbar und $|C| \neq 0$, so ist mit $B = Fu(A)$ auch $CBC^{-1} = Fu(A)$.

Für reguläre Stellen A gibt es nach Definition 3 nur ein $Fu(A)$. Also gilt bei $|C| \neq 0$:

Satz 2a: Für reguläre Stellen A ist $Fu(A)$ mit allen C vertauschbar, die mit A vertauschbar sind.

Bei $|C| = 0$ ist für kleine $u > 0$ sicher $|C + uE| \neq 0$ und $C + uE$ mit A vertauschbar. Aus $(C + uE)Fu(A) = Fu(A) \cdot (C + uE)$ folgt bei $u \rightarrow 0$ die allgemeine Gültigkeit von Satz 2a. (Beispiel 2 zeigt, daß die

entsprechende Behauptung für richtungssinguläre Stellen falsch sein kann.) Da alle $A \in \mathfrak{B}_0$ regulär sind, ist also insbesondere für die in Definition 2 verwendete Folge $A(t): A(t) \cdot F u(A(t)) = F u(A(t)) \cdot A(t)$. Der Grenzübergang zu $t = 0$ liefert

Satz 2b: Für alle $A \in \mathfrak{B}$ ist $F u(A)$ mit A vertauschbar.

Nach diesen einleitenden Sätzen wollen wir die Art der durch eine Matrixfunktion geschaffenen Abhängigkeit klären. Hierzu gehen wir von den $A \in \mathfrak{B}_0$ aus. Es gibt für jedes solches A wenigstens ein C , das A in Diagonalform $CAC^{-1} = (\lambda_\nu)$ transformiert. Nach (2.2) und Satz 2b ist dann $CBC^{-1} = (\mu_\nu)$ mit

$$\mu_\nu = g_\nu(\lambda_\nu; \lambda_1, \dots, \lambda_{\nu-1}, \lambda_{\nu+1}, \dots, \lambda_n),$$

wo g_ν in den Variablen $\lambda_1, \dots, \lambda_{\nu-1}, \lambda_{\nu+1}, \dots, \lambda_n$ symmetrisch sein muß, da man durch geeignete Transformation die entsprechenden Einheitsvektoren vertauschen kann, ohne gemäß (2.2) den Wert von μ_ν zu verändern. g_ν hängt daher von λ_ν und den ersten $(n-1)$ Potenzsummen der $\lambda_\mu \neq \lambda_\nu$ ab. Da man dieselben aber auch als Funktionen von λ_ν und den entsprechenden Potenzsummen aller λ_ν schreiben kann, ist somit

$$\mu_\nu = h(\lambda_\nu; s_1, s_1, \dots, s_{n-1}).$$

Hierbei ist der Index ν bei h weggelassen worden, da wegen der Invarianz gegen Vertauschung aller Grundvektoren die Funktion h für alle ν die gleiche sein muß. h ist eine gewöhnliche stetige Funktion, die durch $F u$ wegen der Unabhängigkeit der Variablen $\lambda_\nu, s_1, \dots, s_{n-1}$ eindeutig bestimmt ist und die Matrixfunktion vermittelt. Ist umgekehrt eine stetige Funktion $h(\lambda; s_1, \dots, s_{n-1})$ vorgegeben, so können wir für alle Matrizen, deren Eigenwerte λ_ν verschieden sind und für welche die Argumente $\lambda_\nu, s_1, \dots, s_{n-1}$ im Inneren des Definitionsbereiches von h liegen, eine Matrix B durch $B = C^{-1} \times \times (h(\lambda_\nu; s_1, \dots, s_{n-1})) \cdot C$ bilden. B ist dann eindeutig bestimmt und hängt stetig von A ab, da bei verschiedenen λ_ν ein stetig verändertes A auch stetige Veränderungen der λ_ν und von C nach sich zieht. Damit sind die Bedingungen von Definition 1 erfüllt.

Es ist daher zweckmäßig, anstelle des allgemeinen Symbols $F u$ die spezielle Gestalt der Abhängigkeit durch die folgende Schreibweise zum Ausdruck zu bringen:

$$(2.4) \quad B = h(A; s_1, \dots, s_{n-1}) \text{ 1. Normaldarstellung.}$$

Durch (2.4) soll gleichzeitig das soeben angegebene Verfahren der Berechnung von B symbolisiert werden. Weiter setzt die 1. Normaldarstellung die Eigenwerte von B in Evidenz.

Haben wir eine Matrix A mit teilweise zusammenfallenden Eigenwerten, die sich jedoch in Diagonalform $CAC^{-1} = (\lambda_\nu)$ bringen läßt („symmetrisierbare Matrix“) und liegen die Argumente $\lambda_\nu, s_1, \dots, s_{n-1}$ im stetigen Definitionsbereich von h , dann können wir jedenfalls wie oben B bilden. A liegt dann in der abgeschlossenen Hülle von $\mathfrak{B}_0$ und B erscheint als Limes von $F u(A(t))$ mit $A(t) \in \mathfrak{B}_0$. Solche A liegen also jedenfalls in $\mathfrak{B}$. Wir wollen uns nun überzeugen, daß jedoch A durchaus eine richtungssinguläre Stelle sein kann, so daß (2.4) nicht den gesamten Wertevorrat von $F u(A)$ für die $A \in \mathfrak{B}$ mit zusammenfallenden Eigenwerten liefert. Dieses Verhalten zeigt

Beispiel 3: Es sei

$$f(1 + r e^{i\varphi}) = \begin{cases} 1 + \sqrt{r} e^{i\frac{\varphi}{2}} & \text{für } |\varphi| \leq \frac{\pi}{2} \\ 1 + \sqrt{r} e^{i\frac{1}{2}(3|\varphi| - \pi) \cdot \text{sign } \varphi} & \text{für } \frac{\pi}{2} \leq |\varphi| \leq \pi. \end{cases}$$

$f(x)$ ist eine für alle x stetige Funktion. Wir setzen $h(\lambda_\nu; s_1, \dots, s_{n-1}) \equiv t(\lambda_\nu)$. Durch direkte Anwendung von (2.4) finden wir bei $A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = E$:

$$Fu(E) = E.$$

Benutzen wir aber die bei $t \rightarrow 0$ gegen E konvergente Folge

$$A(t) = \begin{pmatrix} 1 + it^2 & 0 \\ t\sqrt{2} & 1 - it^2 \end{pmatrix} = V \cdot \begin{pmatrix} 1 + it^2 & 0 \\ 0 & 1 - it^2 \end{pmatrix} \cdot V^{-1} \text{ mit } V = \begin{pmatrix} it\sqrt{2} & 0 \\ 1 & 1 \end{pmatrix},$$

so ist

$$Fu(A(t)) = V \cdot \begin{pmatrix} 1 + te^{i\frac{\pi}{4}} & 0 \\ 0 & 1 + te^{-i\frac{\pi}{4}} \end{pmatrix} V^{-1} = \begin{pmatrix} 1 + te^{i\frac{\pi}{4}} & 0 \\ 1 & 1 + te^{-i\frac{\pi}{4}} \end{pmatrix}$$

und damit bei $t \rightarrow 0$:

$$Fu(E) = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}.$$

Liegt mit A auch $\bar{A}$ in $\mathfrak{B}_0$, so ist bei $A = C^{-1}(\lambda_\nu) C$ und $B = C^{-1}(\mu_\nu) C$ wegen (2.2):

$$Fu(\bar{A}) = Fu(\bar{C}(\lambda_\nu) \bar{C}^{-1}) = \bar{C} Fu(\lambda_\nu) \bar{C}^{-1} = \bar{C}(\mu_\nu) \bar{C}^{-1} = \overline{C^{-1}(\mu_\nu) C} = \bar{B}$$

Wir können daher gegebenenfalls den Bereich $\mathfrak{B}_0$ dadurch erweitern, daß wir alle gespiegelten Matrizen hinzunehmen und dabei $Fu(\bar{A}) = \overline{Fu(A)}$ setzen. Anschließend können wir dann unter Erhaltung dieser Beziehung gemäß Definition 2 den Definitionsbereich $\mathfrak{B}$ bilden, der dann ebenfalls mit jeder Matrix auch ihre Gespiegelte enthält. Es gilt daher

Satz 3: Zu $\mathfrak{B}$ dürfen die Gespiegelten aller Matrizen $A \in \mathfrak{B}$ hinzugenommen werden mit $Fu(\bar{A}) = \overline{Fu(A)}$.

Liegt A in $\mathfrak{B}_0$ und ist $A = \begin{pmatrix} A_1 & 0 \\ 0 & A_2 \end{pmatrix}$, so ist gemäß der durch die erste Normaldarstellung gegebenen Rechenvorschrift $Fu(A) = B = \begin{pmatrix} B_1 & 0 \\ 0 & B_2 \end{pmatrix}$, wobei wir B_1 als eine Matrixfunktion von A_1 auffassen können, die außerdem noch die Invarianten von A_2 als Parameter enthält. Für richtungssinguläre Stellen A zeigt dagegen Beispiel 3, daß aus der Reduziertheit von A nicht die von B folgt. Richtig ist diese Behauptung jedoch, wenn A_1 und A_2 eigenwertfremd sind, da dann die Reduziertheit von B sich aus Satz 2b ergibt. Um nun zu sehen, daß auch in diesem Falle B_1 aus einer Matrixfunktion von A_1 mit den Invarianten von A_2 als Parametern gewonnen wird, beweisen wir zunächst

Hilfssatz 1: Ist gegeben eine Folge $A(t)$ bei $\lim_{t \rightarrow 0} A(t) = A = \begin{pmatrix} A_1 & 0 \\ 0 & A_2 \end{pmatrix}$ mit eigenwertfremden A_1 und A_2 , dann gibt es eine Folge $T(t)$ mit

$$A(t) = T(t) \cdot \begin{pmatrix} A_1(t) & 0 \\ 0 & A_2(t) \end{pmatrix} T^{-1}(t); \lim_{t \rightarrow 0} T(t) = E; \lim_{t \rightarrow 0} A_\nu(t) = A_\nu.$$

Beweis: Der Grad von A_ν sei m_ν : $m_1 + m_2 = n$. Ist nun $\varphi_\nu(y) = 0$ die charakteristische Gleichung von A_ν , so entspricht $\varphi_\nu(y)$ wegen der Eigenwertfremdheit der A_ν für kleine t eindeutig einem Faktor $\varphi_\nu(y, t)$ der charakteristischen Gleichung von $A(t)$. Bilden wir $C_\nu(t) = \varphi_\nu(A(t), t)$, so hat $C_\nu(t)$ genau den Rang $n - m_\nu$. $C_\nu(t)$ bestimmt also vermöge $C_\nu(t) \cdot \xi = 0$ einen m_ν -dimensionalen Unterraum $U_\nu(t)$ als Eigenraum von $A(t)$, der wegen der Konstanz des Ranges bei $t \rightarrow 0$ in den durch $\varphi_\nu(A) \cdot \xi = 0$ definierten Unterraum U_ν übergeht, der durch die ersten m_1 , bzw. letzten m_2 , Grundvektoren e_i aufgespannt wird. Es gibt also in $U_\nu(t)$ m_ν linear unabhängige Vektoren $\tilde{f}_i(t)$ mit $\lim_{t \rightarrow 0} \tilde{f}_i(t) = e_i$. Alle $\tilde{f}_i(t)$ zusammen spannen den ganzen Raum auf. Das gesuchte $T(t)$ ist die Matrix, die e_i in $\tilde{f}_i(t)$ transformiert.

Wir haben nun

Satz 4: Ist $A \in \mathfrak{B}$ und $A = \begin{pmatrix} A_1 & 0 \\ 0 & A_2 \end{pmatrix}$ mit eigenwertfremden A_ν , so ist $Fu(A) = \begin{pmatrix} B_1 & 0 \\ 0 & B_2 \end{pmatrix}$, wo B_1 Matrixfunktion von A_1 mit den Invarianten von A_2 als Parametern ist: $B_1 = Fu(A_1; \{A_2^k\})$.

Beweis: Für $A \in \mathfrak{B}_0$ wurde der Beweis oben geführt. Für die übrigen $A \in \mathfrak{B}$ wird $Fu(A)$ gemäß Definition 2 mit Hilfe einer Folge $A(t)$ gebildet, auf die Hilfssatz 1 anwendbar ist. Es ist dann wegen (2.2):

$$Fu(A(t)) = T(t) \cdot \begin{pmatrix} Fu(A_1(t); \{A_2^k(t)\}) & 0 \\ 0 & Fu(A_2(t); \{A_1^k(t)\}) \end{pmatrix} T^{-1}(t),$$

woraus bei $t \rightarrow 0$ unmittelbar die Behauptung folgt.

In (2.4) kommt die letzte unabhängige Invariante s_n von A nicht mit vor. Man kann sie aber mitnehmen und allgemeiner $B = g(A; s_1, \dots, s_n)$ schreiben. Doch ist diese Darstellung dann nicht mehr eindeutig. Da man umgekehrt s_n als Funktion eines λ_ν und der $s_1, \dots, s_{n-1}$ schreiben kann, läßt sich stets die eindeutige Normaldarstellung (2.4) zurückgewinnen. Man kann nun diesen Sachverhalt ausnutzen, um durch die Einführung von s_n die Abhängigkeit der Funktion g von ihrer ersten Variablen möglichst einfach zu gestalten. Hierzu bestimmen wir für $A \in \mathfrak{B}_0$ die Größen $c_0, c_1, \dots, c_{n-1}$ aus dem Gleichungssystem:

$$(2.5) \quad \sum_{k=0}^{n-1} c_k \cdot \lambda_\nu^k = h(\lambda_\nu; s_1, \dots, s_{n-1}) \quad \nu = 1, 2, \dots, n.$$

Die c_k sind offensichtlich invariant gegen Permutationen der λ_ν und können somit als Funktionen von $s_1, \dots, s_n$ aufgefaßt werden:

$$c_k = c_k(s_1, \dots, s_n).$$

Es wird dann nach der durch (2.4) symbolisierten Rechenvorschrift:

$$B = C^{-1} \cdot \left(\sum_{k=0}^{n-1} c_k \lambda_\nu^k \right) C = \sum_{k=0}^{n-1} c_k \cdot C^{-1}(\lambda_\nu^k) C = \sum_{k=0}^{n-1} c_k \cdot [C^{-1}(\lambda_\nu) C]^k$$

oder

$$(2.6) \quad B = \sum_{k=0}^{n-1} c_k(s_1, \dots, s_n) \cdot A^k. \quad 2. \text{ Normaldarstellung.}$$

$Fu(A)$ kann also stets als ganz-rationale Funktion von A vom Grade $n-1$ aufgefaßt werden, wobei die Koeffizienten Invarianten von A sind. Umgekehrt liefert (2.6) mit beliebigen stetigen Invarianten c_k stets eine stetige Matrixfunktion für einen geeigneten Bereich $\mathfrak{B}_0$.

Die 2. Normaldarstellung gilt zunächst wieder nur für die $A \in \mathfrak{B}_0$. Für andere A können bei Approximation mit Hilfe einer Folge $A(t)$ die gemäß (2.5) gebildeten $c_k(t)$ konvergieren. Dann gehört zwar A zu $\mathfrak{B}$, kann aber durchaus eine richtungssinguläre Stelle sein, so daß (2.6) dann nicht alle Werte von $Fu(A)$ wiedergibt. Dies folgt auch daraus, daß ein gemäß (2.6) gebildetes B automatisch mit allen C vertauschbar ist, die mit A vertauschbar sind.

Da jede Matrix ihrer charakteristischen Gleichung genügt, kann für jedes m die Funktion A^m durch $A^0, A^1, \dots, A^{m-1}$ und $s_1, \dots, s_n$ dargestellt werden. (2.6) ist die Verallgemeinerung dieses Satzes für beliebige Matrixfunktionen.

Wir hatten gesehen, daß wir die erste Normaldarstellung (2.4) auf alle A aus $\mathfrak{B}$ anwenden dürfen, die symmetrisierbar sind und damit jedenfalls wenigstens einen Wert für $Fu(A)$ erhalten. Doch ist die Anwendung von (2.4) wegen der Transformation auf Hauptachsen sehr umständlich. Für alle symmetrisierbaren A läßt sich nun aber (2.4) wie folgt umformen:

Es seien $\lambda_1, \dots, \lambda_m$ mit $m \leq n$ die verschiedenen Eigenwerte von A ; dann bilden wir die m Hilfsmatrizen

$$(2.7) \quad G_k = \prod_{i \neq k} \frac{A - \lambda_i E}{\lambda_k - \lambda_i}, \quad k = 1, \dots, m.$$

Ist nun $CAC^{-1} = (\lambda_\nu)$, so wird CG_kC^{-1} eine Diagonalmatrix, die nur an den Stellen, wo in (λ_ν) der Wert λ_k stand, eine 1 besitzt, sonst aber lauter Nullen hat. Es ist daher gemäß (2.4):

$$C B C^{-1} = \sum_{k=1}^m \mu_k \cdot C G_k C^{-1}$$

und damit

$$(2.8) \quad \begin{cases} h(A; s_1, \dots, s_{n-1}) = \sum_{k=1}^m \mu_k \cdot G_k \text{ mit} \\ \mu_k = h(\lambda_k; s_1, \dots, s_{n-1}) \text{ und} \\ G_k = \prod_{i \neq k} \frac{A - \lambda_i E}{\lambda_k - \lambda_i}. \end{cases} \quad 3. \text{ Normaldarstellung}$$

Die 3. Normaldarstellung ist gültig für alle $A \in \mathfrak{B}_0$ und für alle übrigen symmetrisierbaren $A \in \mathfrak{B}$, für die sie aber evtl. nicht alle Werte liefert. Für nichtsymmetrisierbare A liefert sie falsche Resultate. Sie ist besonders bequem für numerische Rechnungen, vor allem dann, wenn es sich darum handelt, von einer festen Matrix A verschiedene Funktionen zu berechnen, da in diesem Falle die Aufstellung der G_k nur einmal zu geschehen braucht.

§ 3. Gewöhnliche, insbesondere analytische Matrixfunktionen.

Ist die Abhängigkeit $Fu(A)$ durch Anwendung der Potenzreihe einer ganzen analytischen Funktion $f(x)$ auf A gegeben, so ist $\mu_\nu = f(\lambda_\nu)$, so daß in der eindeutig bestimmten ersten Normaldarstellung die explizite

Abhängigkeit von $s_1, \dots, s_{n-1}$ entfällt. Es ist daher sinnvoll, ganz allgemein die Anwendung einer eindeutigen Funktion $f(x)$ auf Matrizen im Sinne der Vorschrift der ersten Normaldarstellung aufzufassen und dann einfach

$$(3.1) \quad B = f(A)$$

zu schreiben. Diese so gewonnene „gewöhnliche Matrixfunktion“ heißt „analytische Matrixfunktion“, wenn $f(x)$ an den Stellen der Eigenwerte der A aus $\mathfrak{B}$ analytisch ist.

Betrachten wir eine Menge von allgemeinen Funktionen für ein und dieselbe Matrix A , so sind die s_k in (2.4) feste Zahlen, so daß der Unterschied zwischen allgemeiner und gewöhnlicher Matrixfunktion verwischt wird. Die Unterscheidung ist jedoch wesentlich, sobald wir eine Matrixfunktion für variabel gedachtes A untersuchen. Im Falle der analytischen Matrixfunktion sind aus der RIEMANNschen Fläche von $f(x)$ geeignete Stücke auszuscheiden, über denen dann $f(x)$ als eindeutige Funktion betrachtet werden kann, wodurch $\mathfrak{B}$ festgelegt wird. Solange A verschiedene Eigenwerte hat, bedeutet die hierdurch geschaffene Einengung des Variabilitätsbereiches von A keine wesentliche Einschränkung, da man ja die Schnitte geeignet legen kann, um alle Zweige von $f(x)$ benutzen zu können. Schwierigkeiten tauchen nur bei denjenigen A auf, bei denen Eigenwerte zusammenfallen, da dann der durch Definition 2 geforderte Grenzübergang dazu führen kann, daß gleichen λ_r verschiedene μ_r entsprechen. (Weiteres hierzu siehe in § 4.) Bei gewöhnlichen Matrixfunktionen ist in Satz 4 für die $A \in \mathfrak{B}_0$ einfach $Fu(A_1; \{A_2^k\}) = f(A_1)$. Satz 4 gestattet daher jetzt die einfachere Formulierung von

$$\text{Satz 5: Ist } A \in \mathfrak{B} \text{ und } A = \begin{pmatrix} A_1 & 0 \\ 0 & A_2 \end{pmatrix} \text{ mit eigenwertfremden } A_r, \text{ so ist } f(A) = \begin{pmatrix} f(A_1) & 0 \\ 0 & f(A_2) \end{pmatrix}.$$

Damit ist die Bestimmung von $f(A)$ auf die A mit lauter gleichen Eigenwerten λ zurückgeführt. Wie Beispiel 3 zeigte, ist Satz 5 ohne die Einschränkung der Eigenwertfremdheit der A_r falsch. Das in Beispiel 3 gezeigte Verhalten ist nun wesentlich dadurch bestimmt, daß $f(x)$ nicht analytisch ist. Wir führen daher jetzt die weitere Untersuchung von $f(A)$ bei A mit lauter gleichen Eigenwerten λ unter der Voraussetzung des analytischen Charakters von $f(x)$ an der Stelle $x = \lambda$ durch. Hierbei ist für die Beurteilung des analytischen Charakters von $f(x)$ auf die Schnitte Rücksicht zu nehmen, durch die $f(x)$ eindeutig gemacht wurde. Hat man z. B. die RIEMANNsche Fläche von $f(x) = \sqrt{x}$ längs der negativ-reellen Achse aufgeschnitten, so zählt diese einschließlich $x = 0$ nicht zu den Argumenten, wo $f(x)$ analytisch ist. Man kann aber durch geschicktes Legen der Schnitte erreichen, daß jedes A erfaßt wird, für welches $f(x)$ im gewöhnlichen Sinne an den Stellen der Eigenwerte von A regulär ist.

Schreiben wir nun:

$$f(x) = g(x - \lambda) + f(\lambda) \quad \text{mit } g(0) = 0,$$

so ist bei $t > 0$ wegen der Verschiedenheit der Eigenwerte $\lambda_r(t)$ der Matrix $A(t)$ aus der gegen A konvergenten Folge:

$$(3.2) \quad f(A(t)) = g(A(t) - \lambda E) + f(\lambda) \cdot E.$$

$g(x)$ habe nun die für kleine x konvergente Entwicklung:

$$g(x) = \sum_{k=1}^{\infty} d_k x^k = \sum_{k=1}^{N-1} d_k x^k + h(x) \quad \text{mit} \quad d_k = \frac{1}{k!} f^{(k)}(\lambda),$$

so daß wir erhalten

$$(3.3) \quad g(A(t) - \lambda E) = \sum_{k=1}^{N-1} d_k \cdot (A(t) - \lambda E)^k + h(A(t) - \lambda E).$$

Gemäß (2.6) ist:

$$h(A(t) - \lambda E) = \sum_{i=1}^{n-1} c_i(t) \cdot (A(t) - \lambda E)^i,$$

wo die $c_i(t)$ bestimmt sind durch (2.5):

$$\sum_{i=0}^{n-1} c_i(t) (\lambda_v(t) - \lambda)^i = h(\lambda_v(t) - \lambda).$$

Bei genügend großem N folgt hieraus wegen $h(x) = O(x^N)$:

$$\lim_{t \rightarrow 0} c_i(t) = 0$$

und damit

$$\lim_{t \rightarrow 0} h(A(t) - \lambda E) = 0.$$

Aus (3.2) und (3.3) erhalten wir nun bei $t \rightarrow 0$ unabhängig von der definierenden Folge $A(t)$ den Wert für $f(A)$, so daß wir wegen $(A - \lambda E)^k = 0$ für $k \geq n$ haben

Satz 6: Ist $f(x)$ analytisch an der Stelle $x = \lambda$ des n -fachen Eigenwertes von A (des Grades n), so ist $A \in \mathfrak{B}$; A ist eine reguläre Stelle der Matrixfunktion mit

$$f(A) = \sum_{k=0}^{n-1} \frac{1}{k!} f^{(k)}(\lambda) \cdot (A - \lambda E)^k.$$

Wenn A reduziert ist, so ist auch $f(A)$ reduziert, so daß also Satz 5 für analytische Matrixfunktionen auch ohne die Einschränkung der Eigenwertfremdheit richtig ist, wenn $f(x)$ an den Stellen der Eigenwerte von A analytisch ist. Es sei nun A' eine Matrix in der reduzierten Gestalt

$$A' = \begin{pmatrix} A_1 & & 0 \\ & A_2 & \\ & & \ddots \\ 0 & & & A_m \end{pmatrix},$$

wo die eigenwertfremden A_ν vom Grade n_ν mit den n_ν -fachen Eigenwerten λ_ν sind. $f(x)$ sei analytisch für $x = \lambda_\nu$. Für jedes ν sind dann die Polynome $\prod_{\mu \neq \nu} (x - \lambda_\mu)^{n_\mu}$ und $(x - \lambda_\nu)^{n_\nu}$ teilerfremd, so daß wir ein Polynom $H_\nu(x)$ bestimmen können durch

$$(3.4) \quad H_\nu(x) \cdot \prod_{\mu \neq \nu} (x - \lambda_\mu)^{n_\mu} \equiv 1 \pmod{(x - \lambda_\nu)^{n_\nu}}.$$

Wir führen nun ein

$$G'_\nu = \dot{H}_\nu(A') \cdot \prod_{\mu \neq \nu} (A' - \lambda_\mu E)^{n_\mu},$$

dann ist

$$(3.5) \quad G'_\nu = \begin{pmatrix} 0 & \dots & 0 & 0 \\ & \ddots & & \\ 0 & & E_\nu & 0 \\ & & & \ddots & \\ & & 0 & & 0 \end{pmatrix} \quad \text{und} \quad G_\nu A' = \begin{pmatrix} 0 & \dots & 0 & 0 \\ & \ddots & & \\ 0 & & A_\nu & 0 \\ & & & \ddots & \\ & & 0 & & 0 \end{pmatrix}.$$

Nach den Sätzen 5 und 6 haben wir nun

$$f(A') = \sum_{\nu=1}^m G'_\nu \cdot \sum_{k=0}^{n_\nu-1} \frac{1}{k!} f^{(k)}(\lambda_\nu) \cdot (A' - \lambda_\nu E)^k.$$

Durch Transformation mit einer beliebigen Matrix C erhalten wir hieraus den

Satz 7: Hat A die Eigenwerte λ_ν mit den Vielfachheiten n_ν , und ist $f(x)$ an den Stellen λ_ν analytisch, so ist $A \in \mathfrak{B}$ mit dem eindeutig bestimmten Werte³⁾

$$f(A) = \sum_{\nu=1}^m G_\nu \cdot \sum_{k=0}^{n_\nu-1} \frac{1}{k!} f^{(k)}(\lambda_\nu) \cdot (A - \lambda_\nu E)^k,$$

wo

$$G_\nu = H_\nu(A) \cdot \prod_{\mu \neq \nu} (A - \lambda_\mu E)^{n_\mu} \text{ ist mit } H_\nu(x) \text{ gemäß (3.4).}$$

Bemerkung 7a: Anstelle von n_ν können die Exponenten $m_\nu \leq n_\nu$ benutzt werden, für die gerade $\prod_{\nu} (A - \lambda_\nu E)^{m_\nu} = 0$ ist; bei symmetrisierbaren A also $m_\nu = 1$. Im letzteren Falle ist dann

$$H_\nu(x) = \frac{1}{\prod_{\mu \neq \nu} (\lambda_\nu - \lambda_\mu)},$$

so daß Satz 7 ein Spezialfall von (2.8) wird⁴⁾.

Bemerkung 7b: Der in Satz 7 angegebene Ausdruck für $f(A)$ kann noch modulo der charakteristischen Gleichung von A reduziert werden, so daß man die Darstellung (2.6) erhält. In der Tat bleiben ja die c_k als Lösung von (2.5) analytische Funktionen der Eigenwerte λ_ν , auch wenn einige derselben zusammenrücken. Sie genügen dann dem aus (2.5) durch Grenzübergang entstehenden Gleichungssystem:

$$(3.6) \quad \sum_{k=r}^{n-1} c_k \binom{k}{r} \lambda_\nu^{k-r} = \frac{1}{r!} f^{(r)}(\lambda_\nu)$$

³⁾ Vgl. FANTAPPIÈ, L.: Le calcul des matrices, C. r. 186, 619–621 (1928), wo durch Anwendung der Theorie der linearen analytischen Funktionalen auf die Komponenten von $f(A)$ gezeigt wird, daß allein von A abhängige Matrizen $H_{\nu k}$ existieren, so daß

$$f(A) = \sum_{\nu=1}^m \sum_{k=0}^{n_\nu-1} H_{\nu k} f^{(k)}(\lambda_\nu)$$

wird.

⁴⁾ Unter der Einschränkung, daß $f(x)$ eine ganze Funktion ist und keine Eigenwerte von A zusammenfallen, bereits zu finden bei A. D. MICHAL, Matrix and Tensor Calculus, S. 18. New York 1947,

bei $r = 0, 1, \dots, n_r - 1$ und $v = 1, 2, \dots, m$; n_r = Vielfachheit des Eigenwertes λ_r .

Aus Satz 7 können wir nun ein Kriterium dafür ableiten, daß bei reellem A auch $f(A)$ reell ist. Von den Eigenwerten von A seien in diesem Falle $\lambda_1, \dots, \lambda_{2r}$ paarweise konjugiert komplex und $\lambda_{2r+1}, \dots, \lambda_m$ reell. Es ist dann nach (3.4): $H_1 = H_2^*, \dots, H_{2r-1} = H_{2r}^*$, während H_{2r+1} bis H_m reelle Polynome sind. Also ist wegen der Realität von A : $G_1 = G_2^*, \dots, G_{2r-1} = G_{2r}^*$, und G_{2r+1} bis G_m sind reelle Matrizen. Notwendig und hinreichend dafür, daß $f(A)$ reell ist, ist also nach Satz 7, wenn wir gemäß Bemerkung 7a die m_r anstelle der n_r benutzen:

$$\sum_{v=1}^m \sum_{k=0}^{m_v-1} \frac{1}{k!} (f^{(k)}(\lambda_v) - [f^{(k)}(\lambda_v^*)]^*) \cdot G_v \cdot (A - \lambda_v E)^k = 0.$$

An (3.5) sehen wir nun, daß die Matrizen $G_v \cdot (A - \lambda_v E)^k$ mit $0 \leq k \leq m_v - 1$ voneinander linear unabhängig sind. Damit haben wir den

Satz 8: Notwendig und hinreichend, daß bei reellen A die analytische Matrixfunktion $f(A)$ ebenfalls reell ist, ist

$$f^{(k)}(\lambda_v) = [f^{(k)}(\lambda_v^*)]^* \text{ für } k = 0, 1, \dots, m_v - 1,$$

wenn $\prod_v (x - \lambda_v)^{m_v}$ das Polynom niedrigsten Grades ist, welches A annulliert.

Bemerkung 8a: Ist n_r die Vielfachheit von λ_r , so ist $m_r \leq n_r$, so daß die Bedingung von Satz 8 für $k = 0, 1, \dots, n_r - 1$ jedenfalls hinreichend ist. Bei symmetrisierbarem A ist $m_r = 1$, so daß nur $f(\lambda_r) = (f(\lambda_r^*))^*$ zu fordern ist.

Bemerkung 8b: Ist $f(x)$ eine reelle analytische Funktion, d. h. ist $f(x)$ reell auf einem Stück der reellen Achse, und wird derjenige Zweig von $f(x)$ gewählt, für den $f(x^*) = (f(x))^*$ ist, so sagt Satz 8 aus, daß $f(A)$ genau dann reell ist, wenn $f(\lambda_v)$ reell ist für die reellen unter den λ_v . So ist z. B. in A reell bei den reellen A , deren reelle Eigenwerte (falls vorhanden) positiv sind. Satz 8 zeigt aber, daß die Möglichkeit der Realität von $f(A)$ im allgemeinen größer ist, da die Bedingung von Satz 8 für bestimmte λ_v ganz zufällig bis zu $k = m_v - 1$ erfüllt sein kann, während die höheren Ableitungen von $f(x)$ an dieser Stelle nicht mehr dieser Bedingung zu genügen brauchen. Dies hätte zur Folge, daß für jedes reelle D dann $f(A + tD)$ auch für kleines $t > 0$ nicht mehr reell ist, so daß man A in Übertragung einer bei den reellen Funktionen üblichen Sprechweise als „isolierte Matrix für $f(A)$ “ anzusprechen hat.

§ 4. Die Umkehrung analytischer Matrixfunktionen.

Wir hatten in § 3 für die Definition der analytischen Matrixfunktion eine Aufschneidung der RIEMANNSchen Fläche von $f(x)$ vorgeschrieben, um zu einer eindeutigen Matrixfunktion zu gelangen. Man wird nun die Frage aufwerfen, welches die Gesamtheit aller $f(A)$ ist, wenn wir bei der gemäß Definition 2 zu erfolgenden Bestimmung von $f(A)$ mit Hilfe einer Folge $A(t)$ die volle Bewegungsfreiheit der Eigenwerte von $A(t)$ auf der RIEMANNSchen Fläche zulassen. Solange die Eigenwerte von A alle untereinander und von den singulären Stellen von $f(x)$ verschieden sind, kann sich natürlich nichts Neues ergeben, da wir durch geeignete Aufschneidung stets

erreichen können, daß $A \in \mathfrak{B}_0$ ist. Wenn dagegen A teilweise zusammenfallende Eigenwerte hat und $\lim_{t \rightarrow 0} A(t) = A$ ist, wo $A(t)$ für $t > 0$ die verschiedenen

Eigenwerte $\lambda_\nu(t)$ hat, so bilden die $\lambda_\nu(t)$ „Bahnkurven“ in der x -Ebene mit teilweise zusammenfallenden Endpunkten. Wir können dann eine Aufschneidung der RIEMANNSchen Fläche so bilden, daß die Schnittkurven diese Bahnlinien nicht treffen. Bei Bahnkurven mit gleichen Endpunkten λ muß der Schnitt dabei natürlich durch λ geführt werden. Auf diese Weise erhalten wir eine durch diese Aufschneidung festgelegte Matrixfunktion, für die A eine Randstelle wird. Diese Betrachtung zeigt, daß die Gesamtheit aller möglichen $f(A)$ bei freier Bewegungsmöglichkeit auf der RIEMANNSchen Fläche identisch wird mit der Gesamtheit aller Werte von $f(A)$ bei allen möglichen Aufschneidungen der RIEMANNSchen Fläche. Unsere Einschränkung ist damit als nicht wesentlich für die Ermittlung des gesamten Wertevorrates einer mehrdeutigen analytischen Matrixfunktion erkannt.

Ist $y = f(x)$ die Umkehrfunktion einer eindeutigen analytischen Funktion $x = \varphi(y)$, so gilt bei $B(t) = f(A(t))$ für die definierende Folge $A(t)$ wegen der Verschiedenheit der Eigenwerte von $A(t)$: $A(t) = \varphi(B(t))$ und damit $A = \varphi(B)$. B ist also Lösung der Matrixgleichung

$$(4.1) \quad A = \varphi(B)$$

mit bekanntem A . Umgekehrt liefert jede Lösung B von (4.1) einen der möglichen Werte von $f(A)$ bei passender Aufschneidung. Für solche $f(x)$ ist also die Aufgabe der Gewinnung aller Werte von $f(A)$ äquivalent zur Aufgabe der Lösung der Matrixgleichung (4.1), in der B eine reguläre Stelle für die Matrixfunktion $\varphi(B)$ darstellt. Aus dieser Äquivalenz folgt in Verbindung mit Satz 5 weiterhin, daß wir uns bei der Lösung von (4.1) auf die A mit lauter zusammenfallenden Eigenwerten beschränken können. Außerdem können wir gemäß § 2 voraussetzen, daß sich A in der Normalform (2.3) befindet, wo jetzt alle λ_k gleich λ sind.

Zu jeder Umordnung der Kästchen von A gehöre nun eine bestimmte Matrix U , die diese Umordnung bewirkt. T bezeichne eine mit A vertauschbare Matrix.

Denken wir uns nun das unbekannte B vermöge der Transformation mit einem V ebenfalls in Normalform (2.3) gebracht,

$$B = V \cdot \begin{pmatrix} B_1 & & 0 \\ & \ddots & \\ 0 & & B_k \end{pmatrix} V^{-1},$$

wo die Eigenwerte μ_ν der B_ν teilweise gleich sein können, jedoch stets $\varphi(\mu_\nu) = \lambda$ gilt, so ist nach den Ergebnissen des § 3:

$$(4.2) \quad A = V \cdot \begin{pmatrix} \varphi(B_1) & & 0 \\ & \ddots & \\ 0 & & \varphi(B_k) \end{pmatrix} V^{-1}$$

Ist nun beliebig gewählt

$$W = \begin{pmatrix} W_1 & & 0 \\ & \ddots & \\ 0 & & W_k \end{pmatrix},$$

so daß $W^{-1} \varphi(B_\nu) W_\nu$ in Normalform (2.3) gebracht wird, so ist $VW = TU$ mit passenden T und U und damit

$$\begin{aligned}
 A &= V \cdot W \begin{pmatrix} W_1^{-1} \varphi(B_1) W_1 & & 0 \\ & \ddots & \\ 0 & & W_k^{-1} \varphi(B_k) W_k \end{pmatrix} W^{-1} V^{-1} = \\
 &= T U \begin{pmatrix} \varphi(B'_1) & & 0 \\ & \ddots & \\ 0 & & \varphi(B'_k) \end{pmatrix} U^{-1} T^{-1}
 \end{aligned}$$

mit

$$B'_v = W_v^{-1} B_v W_v.$$

Wir haben also wegen $A T = T A$

$$(4.3) \quad \left\{ \begin{array}{l} U^{-1} A U = \varphi \begin{pmatrix} B & & 0 \\ & \ddots & \\ 0 & & B'_k \end{pmatrix} \\ \text{und} \\ B = T \cdot U \cdot \begin{pmatrix} B'_1 & & 0 \\ & \ddots & \\ 0 & & B'_k \end{pmatrix} U^{-1} T^{-1}. \end{array} \right.$$

Hierbei sind die B'_v -Matrizen, deren Normalform (2.3) nur aus einem einzigen Kästchen besteht: irreduzible Matrizen.

Zur Auffindung von B wird also A zunächst durch ein U^{-1} umgeordnet, die Kästchen nach der Umordnung dann in neue Kästchen A'_v zusammengefaßt, die als $\varphi(B'_v)$ mit irreduziblen B'_v geschrieben werden können. B erhält man dann durch Rücktransformationen der aus den B'_v bestehenden Matrix mit Hilfe von U und zusätzliche Transformation durch eine beliebige mit A vertauschbare Matrix T . Dabei ist bis jetzt noch nicht klar, ob zur Auffindung aller B auch alle der endlich vielen U und der endlich vielen Zusammenfassungen zu den A'_v statthaft sind. Dagegen ist T sicher völlig beliebig wählbar, da bei $A = \varphi(B)$ auch $A = T A T^{-1} = T \varphi(B) T^{-1} = \varphi(T B T^{-1})$ ist, so daß mit B auch stets $T B T^{-1}$ eine Lösung darstellt. Es ist durchaus möglich, daß es zu einem der A'_v mehrere zugehörige B'_v gibt: $B'_{v1}, B'_{v2}, \dots$ mit den jeweils zusammenfallenden Eigenwerten $\mu_{v1}, \mu_{v2}, \dots$. Ist hierbei $\mu_{vi} = \mu_{vk}$, so sind B'_{vi} und B'_{vk} wegen der Irreduzibilität transformierte voneinander: $B'_{vi} = S B'_{vk} S^{-1}$. Es ist dann $A'_v = \varphi(B'_{vi}) = S \varphi(B'_{vk}) S^{-1} = S A'_k S^{-1}$. A'_v ist also mit S vertauschbar. Wir können daher annehmen, daß S bereits in T enthalten ist. Die wesentlich verschiedenen Möglichkeiten für B'_v müssen also zu lauter verschiedenen Werten für $\mu_v = f(\lambda)$ gehören. Hieraus folgt

Satz 9: Ist T eine mit A vertauschbare Matrix mit $|T| \neq 0$, so ist mit B auch $T B T^{-1}$ eine Lösung der Matrixgleichung $A = \varphi(B)$ mit eindeutigem analytischen $x = \varphi(y)$. Bis auf Transformation mit allen T ist die Lösung der Matrixgleichung höchstens endlich- oder abzählbar unendlich-deutig, je nachdem dies die Umkehrfunktion $y = f(x)$ von $x = \varphi(y)$ ist.

Es bleibt jetzt nur noch die Frage, wie ein A' aussehen muß, damit es als $\varphi(B')$ mit irreduziblem B' geschrieben werden kann. Dies läßt sich nun leicht klären unter Verwendung der beiden folgenden Hilfssätze:

Hilfssatz 2: Es bezeichne Q_n die Matrix vom Grade n mit den Komponenten

$$q_{ik} = \begin{cases} 1 & \text{für } k - i = 1 \\ 0 & \text{für } k - i \neq 1 \end{cases}.$$

Dann ist bei $a_l \neq 0$ mit $l \geq 1$ die Matrix

$$P = a_l \cdot Q_n^l + a_{l+1} \cdot Q_n^{l+1} + \dots + a_{n-1} \cdot Q_n^{n-1}$$

äquivalent zu Q_n^l .

Beweis: Es seien $e_1, \dots, e_n$ die Grundvektoren. Führen wir noch $e_r = 0$ für $r \leq 0$ ein, so gilt allgemein $Q_n^l e_r = e_{r-l}$ und damit:

$$(4.4) \quad P e_r = \sum_{\kappa=0}^{\infty} a_{l+\kappa} e_{r-l-\kappa}.$$

Aus der Funktion $\psi(x) = a_l + a_{l+1}x + \dots$ gewinnen wir für $r = 0, 1, \dots$ Koeffizienten $b_{r,\mu}$ gemäß

$$\psi^{-r}(x) = \sum_{\mu=0}^{\infty} b_{r,\mu} x^{\mu}.$$

Wir bilden nun die Vektoren

$$\tilde{f}_s = \sum_{\mu=0}^{\infty} b_{r,\mu} e_{s-\mu} \quad (s = 1, \dots, n),$$

wobei r so bestimmt wird, daß $0 < s - r l \leq l$ ist. Die $\tilde{f}_s$ sind dann wegen $b_{r,0} \neq 0$ linear unabhängig voneinander, und es gilt wegen (4.4):

- a) für $r = 0$, also $s \leq l$: $P \cdot \tilde{f}_s = 0$;
- b) für $r \geq 1$:

$$\begin{aligned} P \tilde{f}_s &= \sum_{\mu=0}^{\infty} \sum_{\kappa=0}^{\infty} b_{r,\mu} a_{l+\kappa} e_{s-\mu-l-\kappa} = \sum_{\omega=0}^{\infty} e_{s-l-\omega} \cdot \sum_{\mu+\kappa=\omega} b_{r,\mu} a_{l+\kappa} = \\ &= \sum_{\omega=0}^{\infty} b_{r-1,\omega} e_{s-l-\omega} = \tilde{f}_{s-l}. \end{aligned}$$

Wählen wir nun die $\tilde{f}_s$ als neue Grundvektoren, so wird P in Q_n^l transformiert.

Hilfssatz 3: Die Normalform (2.3) von Q_n^l besteht aus k Kästchen vom Grade $(m+1)$ und $(l-k)$ Kästchen vom Grade m , wenn $n = m l + k$ mit $k < l$ ist.

Beweis: Es ist $Q_n^l e_r = e_{r-l}$, wenn wieder $e_0 = e_{-1} = \dots = 0$ gesetzt ist. Wählen wir als Transformierende die Matrix, die die e_r in die folgende Reihenfolge bringt:

$$\left. \begin{array}{l} e_1, e_{1+l}, e_{1+2l}, \dots \\ e_2, e_{2+l}, \dots \\ \dots \dots \dots \\ e_k, e_{k+l}, \dots \end{array} \right\} \text{ je } (m+1) \text{ Stück}$$

$$\left. \begin{array}{l} e_{k+1}, e_{k+1+l}, \dots \\ \dots \dots \dots \\ e_l, e_{2l}, \dots \end{array} \right\} \text{ je } m \text{ Stück}$$

so wird Q_n in die angegebene Normalform transformiert.

Bemerkung: Die Hilfssätze 2 und 3 gelten trivialerweise auch bei $l \geq n$, da dann $Q_n^l = 0$ ist; nicht dagegen bei $l = 0$.

Es möge nun $A' = \varphi(B')$ mit irreduziblem B' vom Grade n sein. B' ist also äquivalent zu $\mu E + Q_n$, wo Q_n wie in Hilfssatz 2 definiert ist. Es ist dann A' äquivalent zu $\varphi(\mu E + Q_n)$. Hat nun $\varphi(y)$ an der Stelle μ die Entwicklung

$$\varphi(y) = \lambda + a_l(y - \mu)^l + a_{l+1}(y - \mu)^{l+1} + \dots \text{ mit } a_l \neq 0 \text{ bei } l \geq 1,$$

so wird nach Satz 6:

$$\varphi(\mu E + Q_n) = \lambda E + a_l Q_n^l + a_{l+1} Q_n^{l+1} + \dots,$$

was nach den Hilfssätzen 2 und 3 in der Normalform (2.3) k Kästchen vom Grade $(m+1)$ und $(l-k)$ Kästchen vom Grade m hat, wenn $n = ml + k$ mit $k < l$ ist. Dies ist dann auch die Normalform von A' . Hat umgekehrt A' diese Normalform, so geht es gemäß Hilfssatz 3 bei einer Umordnung der Zeilen und Spalten über in $\lambda E + Q_n$. Ist nun $y = f(x) = \mu + b \sqrt[l]{x - \lambda} + \dots$, so wird, abgesehen von dieser Umordnung, $B' = f(A') = \mu E + b \cdot Q_n + \dots$ die bis auf Transformation mit einer mit A' vertauschbaren Matrix T' eindeutig bestimmte Lösung der Matrixgleichung $A' = \varphi(B')$ mit irreduziblem B' des Eigenwertes μ . Damit haben wir einen vollkommenen Überblick über die Lösungsmannigfaltigkeit von $B = f(A)$ bei solchen A mit lauter gleichen Eigenwerten λ gewonnen:

Satz 10: A möge nur den mehrfachen Eigenwert λ haben. Es befinde sich in Normalform. Es seien

$$f_\nu(x) = \mu_\nu + b_\nu \cdot \sqrt[l_\nu]{x - \lambda} + \dots, \quad \nu = 1, 2, \dots$$

die verschiedenen Zweige von $f(x)$ an der Stelle $x = \lambda$. Dann ordne man die Kästchen von A mit Hilfe eines U^{-1} so um, daß man mit geeigneten Zahlen k_ν und m_ν bei $0 \leq k_\nu < l_\nu$ und $m_\nu \geq 0$ Zusammenfassungen A'_ν von k_ν Kästchen vom Grade $(m_\nu + 1)$ mit $(l_\nu - k_\nu)$ Kästchen vom Grade m_ν bilden kann. Es ist dann $f(A'_\nu) = B'_\nu$ mit irreduziblem B'_ν , wenn der Zweig f_ν verwendet wird. $f(A)$ entsteht dann aus der aus den B_ν gebildeten Matrix durch Rücktransformation mit U und nachträglicher Transformation mit einer beliebigen mit A vertauschbaren Matrix T .

Bemerkung 10a: Sind insbesondere alle $l_\nu = 1$, so wird $f(x)$ auf die Kästchen der Normaldarstellung von A einzeln angewendet und liefert für jedes Kästchen so viele Möglichkeiten, wie Zweige von $f(x)$ vorhanden sind. Die Anwendung von U ist in diesem Falle unnötig. $f(A)$ ist stets lösbar.

Bemerkung 10b: Ist für ein bestimmtes A die angegebene Vorschrift nicht durchführbar, so ist $B = f(A)$ als Umkehrung von $A = \varphi(B)$ unlösbar.

In dem eingangs genannten Beispiel $f(x) = \sqrt{x}$ ist für $\lambda \neq 0$: $l_1 = l_2 = 1$. Für ein Kästchen von A der Gestalt $\lambda E + Q$, E und Q vom Grade n , ist:

$$f(\lambda E + Q) = \pm \sqrt[n]{\lambda} \cdot \sum_{\alpha=0}^{n-1} \left(\frac{1}{\alpha} \right) \left(\frac{Q}{\lambda} \right)^\alpha$$

genau zweideutig. Hat A die Zahl von z_s Kästchen vom Grade s , so hat $f(A)$ bis auf Äquivalenz genau $II(z_s + 1)$ verschiedene Lösungen. Ist dagegen $\lambda = 0$, so ist $l = 2$. Eine Zusammenfassung A_s besteht dann (Typus I)

entweder aus 2 Kästchen des gleichen Grades $m_s \geq 1$; es ist dann nach Hilfsatz 2 $\sqrt{A_s}$ äquivalent zu Q_{2m_s} . Oder A'_s besteht (Typus II) aus 2 Kästchen der Grade $m_s + 1$ und $m_s \geq 1$; es ist dann $\sqrt{A'_s}$ äquivalent zu Q_{2m_s+1} . Oder A'_s ist (Typus III) ein Kästchen vom Grade 1 mit $\sqrt{A'_s} = 0$. Nach vorgenommener Zusammenfassung ist hier $f(A)$ bis auf Transformation durch ein mit A vertauschbares T eindeutig bestimmt. Bis auf Äquivalenz ist die Anzahl der Lösungen von $f(A)$ also gleich der Anzahl der Möglichkeiten, die Matrix A in α_s Stück des Typus I mit $m_s = r$, β_s Stück des Typus II mit $m_s = r$ und β'_s Stück des Typus III zusammenzufassen. Hat nun A wieder z_s Kästchen vom Grade s , so ist:

$$(4.5) \quad z_s = 2\alpha_s + \beta_{s-1} + \beta_s \quad \text{für } s = 1, 2, \dots, N,$$

wo N der größte vorkommende Grad eines Kästchens von A' ist. Dabei ist $\beta_N = 0$ zu setzen. Um die möglichen A'_s zu finden, haben wir (4.5) bei gegebenen z_s mit nichtnegativen α_s und β_s zu lösen. Nehmen wir an, wir hätten eine solche Lösung, so können wir, falls vorhanden, ein $\beta_s \geq 2$ um 2 vermindern und dafür α_s und α_{s+1} um 1 erhöhen und so eine neue Lösung erhalten. Wenn überhaupt, so gibt es also auch eine Lösung α'_s, β'_s , wo alle $\beta'_s = 0$ oder $= 1$ sind. Aus (4.5) folgt unmittelbar, daß dabei

$$\beta'_s = z_{s+1} + z_{s+2} + \dots + z_N \pmod{2}$$

ist, woraus sich dann $\alpha'_s = \frac{1}{2}(z_s - \beta_{s-1} - \beta'_s)$ ergibt, was automatisch ganzzahlig ist. Ist dabei ein $\alpha'_s < 0$, so ist $\sqrt{A}$ nicht lösbar. Sind dagegen alle $\alpha'_s \geq 0$, so gibt es wenigstens diese eine Lösung, aus der wir alle anderen durch mehrmalige Anwendung des folgenden Prozesses gewinnen: Zwei aufeinanderfolgende $\alpha_s \geq 1$ und $\alpha_{s+1} \geq 1$ werden beide um 1 vermindert und dafür β_s um 1 erhöht. Die Anzahl der Lösungen von (4.5) ist dann gleich der Anzahl der nichtäquivalenten Lösungen von $f(A) = \sqrt{A}$. Man bemerke übrigens, daß es stets wenigstens eine Lösung gibt, wenn alle $z_s \geq 1$ sind, da $\beta'_{s-1} + \beta'_s = z_s \pmod{2}$ und $\beta'_{s-1} + \beta'_s \leq 2$ und damit sicher $z_s - \beta'_{s-1} - \beta'_s \geq 0$ ist.

Für zweireihige Matrizen zeigen diese Ergebnisse die folgenden Möglichkeiten:

$$\lambda \neq 0: \quad \sqrt{\begin{pmatrix} \lambda & 0 \\ 0 & \lambda \end{pmatrix}} = \pm \sqrt{\lambda} \cdot E, \quad \sqrt{\lambda} \cdot \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \quad 3 \text{ Lösungen}$$

$$\sqrt{\begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix}} = \pm \sqrt{\lambda} \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \quad 2 \text{ Lösungen}$$

$$\lambda = 0: \quad \sqrt{\begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}} \text{ mit } \begin{cases} \alpha_1 = 0, \beta_0 = 2: \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \\ \alpha_1 = 1, \beta_0 = 0: \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \end{cases} \quad 2 \text{ Lösungen}$$

$$\sqrt{\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}} \text{ mit } \begin{cases} z_1 = 0, z_2 = 1; \\ \text{also } \beta'_0 = \beta'_1 = 1; \\ \alpha'_1 = -1 \end{cases} \quad \text{keine Lösung.}$$

§ 5. Differentiale analytischer Matrixfunktionen.

Es sei jetzt wieder $f(A)$ eine eindeutige analytische Matrixfunktion im Sinne von § 3. A sei eine Matrix, für deren Eigenwerte $f(x)$ bei passender Wahl der Aufschneidung der RIEMANNschen Fläche analytisch ist. Nach Satz 7 ist dann bei $t > 0$:

$$(5.1) \quad \lim_{t \rightarrow 0} f(A + tD) = f(A)$$

für beliebiges D ; insbesondere braucht nicht darauf geachtet zu werden, ob $A + tD$ zusammenfallende Eigenwerte besitzt oder nicht. Es hat daher einen eindeutigen Sinn, die „Variation in der Richtung D “ einzuführen⁵⁾ gemäß

$$(5.2) \quad \delta f(A) = f(A + tD) - f(A), \quad t > 0, \quad t \ll 1.$$

Nach obiger Voraussetzung existiert auch eindeutig $f'(A)$. Es liegt daher nahe zu fragen, ob man $\delta f(A)$ bis auf höhere Potenzen in t durch $t \cdot f'(A) D$ ersetzen darf. Daß dies nicht allgemein möglich ist, zeigt bereits das Beispiel $f(x) = x^2$, wo

$$(5.3a) \quad \delta f(A) - t f'(A) D = t \cdot (DA - AD) + O(t^2)$$

ist, so daß die Vertauschbarkeit von D mit A zu verlangen wäre. Ist jedoch nun G eine beliebige mit A vertauschbare Matrix, so folgt aus (5.3a) nach Multiplikation mit G und Spurbildung wegen der für Spuren gültigen Rechenregel $\{F_1 F_2 \dots F_n\} = \{F_n F_1 \dots F_{n-1}\}$:

$$(5.3b) \quad \{G(\delta f(A) - t f'(A) D)\} = O(t^2) \quad \text{auch bei } AD \neq DA.$$

Die Gültigkeit einer solchen Beziehung unabhängig von D ist deshalb besonders wichtig, weil in den physikalischen Anwendungen die Invarianten der vorkommenden Tensoren wesentlich sind, die sich ja durch die Spuren ausdrücken lassen. Es ist daher bemerkenswert, daß wir den soeben in einem ganz speziellen Beispiel gefundenen Sachverhalt allgemein aussprechen können im

Satz II: Ist $f(A)$ eine analytische Matrixfunktion und $f(x)$ an den Stellen der Eigenwerte von A regulär, so gilt bei $\delta f(A) = f(A + tD) - f(A)$:

$$\begin{aligned} \delta f(A) &= t \cdot f'(A) D + O(t^2), & \text{falls } AD = DA \\ \{G(\delta f(A) - t \cdot f'(A) D)\} &= O(t^2), & \text{falls } AG = GA \text{ und } D \text{ beliebig.} \end{aligned}$$

Wir beweisen den Satz zunächst für den Spezialfall, daß A lauter gleiche Eigenwerte λ hat und führen dann den allgemeinen Fall darauf zurück. Setzen wir $f(x) = g(x - \lambda)$, so wird

$$\delta f(A) = g(A - \lambda E + tD) - g(A - \lambda E) = \delta g(A - \lambda E),$$

$$f'(A) = g'(A - \lambda E)$$

$$\text{und} \quad G \cdot (A - \lambda E) = (A - \lambda E) \cdot G.$$

⁵⁾ Auf solche Variationen wird man in den physikalischen Anwendungen zwangsläufig geführt, z. B. wenn man die Änderung des Spannungstensors bei zeitlich variablem Verzerrungszustand studieren will. Vgl. RICHTER, H.: Das isotrope Elastizitätsgesetz. Z. ang. Math. Mech. 28, 205—209 (1948).

Wir können daher ohne Einschränkung der Allgemeinheit $\lambda = 0$ annehmen. Es sei nun für kleine x :

$$f(x) = \sum_{\nu=0}^{\infty} a_{\nu} x^{\nu}.$$

Wir setzen dann mit noch zu bestimmendem N :

$$f_1(x) = \sum_{\nu=0}^{N-1} a_{\nu} x^{\nu}, \quad f_2(x) = \sum_{\nu=N}^{\infty} a_{\nu} x^{\nu};$$

dann ist

$$\delta f(A) - t f'(A) D = [\delta f_1(A) - t f'_1(A) D] + [\delta f_2(A) - t f'_2(A) D],$$

so daß es genügt, den Satz einzeln für $f_1(x)$ und $f_2(x)$ zu beweisen. Wir haben nun

$$\begin{aligned} \delta f_1(A) - t f'_1(A) D &= \\ &= t \cdot \sum_{\nu=0}^{N-1} a_{\nu} (A^{\nu-1} D + A^{\nu-2} D A + \dots + D A^{\nu-1} - \nu A^{\nu-1} D) + O(t^2), \end{aligned}$$

woraus die Behauptung des Satzes für $f_1(A)$ unmittelbar folgt.

Schreiben wir weiter gemäß (2.6)

$$f_2(A + tD) = \sum_{k=0}^{n-1} c_k(t) (A + tD)^k \quad \text{und} \quad f'_2(A + tD) = \sum_{k=0}^{n-1} d_k(t) (A + tD)^k,$$

so erhält man aus (3.6) bei genügend großem N wegen $f_2(x) = O(x^N)$:

$$c_k(t) = O(t^2) \quad \text{und} \quad d_k(t) = O(t^2)$$

und damit

$$\delta f_2(A) - t f'_2(A) \cdot D = O(t^2),$$

was die Richtigkeit der Behauptungen für $f_2(A)$ zeigt.

Es möge nun A teilweise verschiedene Eigenwerte haben. Nach einer geeigneten Transformation können wir annehmen, daß A reduziert ist:

$$A = \begin{pmatrix} A_1 & 0 \\ 0 & A_2 \end{pmatrix}$$

mit eigenwertfremden A_{ν} . Eine beliebige Matrix P des gleichen Grades wie A schreiben wir dann in der Gestalt

$$P = \begin{pmatrix} P_{11} & P_{12} \\ P_{21} & P_{22} \end{pmatrix},$$

wo P_{11} vom gleichen Grad wie A_1 ist, und nennen:

$$(5.4) \quad \begin{cases} P \text{ reduziert:} & P = (R), \text{ wenn } P_{12} = P_{21} = 0, \\ P \text{ komplementär:} & P = (K), \text{ wenn } P_{11} = P_{22} = 0. \end{cases}$$

Es gelten dann die Rechenregeln:

$$(5.5) \quad \begin{cases} (R) \cdot (R) = (R); & (K) \cdot (K) = (K) \\ (K) \cdot (R) = (K); & (R) \cdot (K) = (K) \\ \{K\} = 0. \end{cases}$$

Das gegebene D schreiben wir als Summe:

$$(5.6) \quad D = D_r + D_k \quad \text{mit} \quad D_r = (R) \quad \text{und} \quad D_k = (K).$$

Schließlich ist wegen der Eigenwertfremdheit der A_r und der Vertauschbarkeit von G mit A :

$$(5.7) \quad G = (R).$$

Ist D mit A vertauschbar, so ist $D_k = 0$. Durch Satz 5 wird unsere Behauptung auf die entsprechende Behauptung für die A_r und damit schließlich auf die für A mit lauter gleichen Eigenwerten zurückgeführt. Sei nun $D_k \neq 0$, so schreiben wir für den Beweis der zweiten Behauptung unseres Satzes gemäß (2.6) für kleine t und u unter Beachtung von (5.5):

$$\begin{aligned} f(A + tD_r + uD_k) &= \sum_{\nu=1}^{n-1} c_\nu(t, u) (A + tD_r + uD_k)^\nu = \\ &= \sum_{\nu=1}^{n-1} c_\nu(t, u) [(A + tD_r)^\nu + (K) + O(u^2)], \end{aligned}$$

oder

$$(5.8) \quad \begin{aligned} f(A + tD_r + uD_k) &= f(A + tD_r) + \\ &+ \sum_{\nu=1}^{n-1} [c_\nu(t, u) - c_\nu(t, 0)] \cdot (A + tD_r)^\nu + (K) + O(u^2). \end{aligned}$$

Bei festem t hängen die $c_\nu(t, u)$ eindeutig analytisch ab von den Potenzsummen

$$s_x(u) = \{(A + tD_r + uD_k)^x\} = \{(A + tD_r)^x + (K) + O(u^2)\} = s_x(0) + O(u^2).$$

Also ist $c_\nu(t, u) - c_\nu(t, 0) = O(u^2)$, so daß aus (5.8) folgt, wenn wir schließlich speziell $u = t$ setzen;

$$f(A + tD) = f(A + tD_r) + (K) + O(t^2).$$

Wir erhalten dann unter Beachtung von $f'(A) = (R)$:

$$\begin{aligned} \delta f(A) - t \cdot f'(A) D &= f(A + tD) - f(A) - t \cdot f'(A) D = \\ &= f(A + tD_r) - f(A) - t \cdot f'(A) D_r + (K) + O(t^2), \end{aligned}$$

und da für $D = D_r$ die erste Behauptung des Satzes bereits bewiesen ist;

$$\delta f(A) - t \cdot f'(A) D = (K) + O(t^2),$$

woraus nach Multiplikation mit G und Spurbildung unter Beachtung von (5.7) und (5.5) die Behauptung folgt.

(Eingegangen am 13. Mai 1949)