COMPOSITIO MATHEMATICA

A. HERRMANN

Über Matrixgleichungen und die Zerlegung von Polynomen in Linearfaktoren

Compositio Mathematica, tome 1 (1935), p. 284-302

<http://www.numdam.org/item?id=CM_1935__1__284_0>

© Foundation Compositio Mathematica, 1935, tous droits réservés.

L'accès aux archives de la revue « Compositio Mathematica » (http://http://www.compositio.nl/) implique l'accord avec les conditions générales d'utilisation (http://www.numdam.org/conditions). Toute utilisation commerciale ou impression systématique est constitutive d'une infraction pénale. Toute copie ou impression de ce fichier doit contenir la présente mention de copyright.

NUMDAM

Article numérisé dans le cadre du programme
Numérisation de documents anciens mathématiques
http://www.numdam.org/

Über Matrixgleichungen und die Zerlegung von Polynomen in Linearfaktoren

von

A. Herrmann

Köthen i. Anh.

Einleitung. In einer Reihe von Arbeiten [1]) sind Matrixgleichungen vom Typus

$$(1) \qquad P(X; A) = X^t + a_1 X^{t-1} + \ldots + a_t; \qquad a_i = r_i(A)$$

untersucht worden, wobei die Koeffizienten a_i rationale Funktionen einer vorgegebenen Matrix $A = (a_{ik})$ bedeuten, eine etwa vorhandene Lösung X erscheint dabei selbst als rationale Funktion $\xi(A)$; weitere Methoden liefern noch allgemeinere Lösungen. Wir werden im Folgenden zunächst im Ringe $R(A)$ der rationalen Funktionen der Matrix A Lösungen aufsuchen und ihre Anzahl bestimmen, und zwar in der Weise, daß wir zu einem zu $R(A)$ isomorphen Ring übergehen, in dem sich dann die weitere Untersuchung sehr einfach gestalten wird. Die Beweise der bekannten Sätze werden sich dann ganz elementar mit Kongruenzbetrachtungen führen lassen. Weiter werden wir erkennen, daß sich unter den Lösungen Systeme von „zueinander konjugierten" Matrizen herausgreifen lassen, denen eine Zerfällung des Polynoms $P(X, A)$ in t Linearfaktoren entspricht. Nachdem wir die

---

[1]) H. Kreis [Vierteljschr. Naturforsch. Ges. Zürich **53** (1908), 366—376].

W. E. Roth, On the matric equation $P(X) = A$ [Trans. Amer. Math. Soc. **30** (1928), 579—596].

W. E. Roth, On the equation $P(X, A) = 0$ in matrices [ebenda **35** (1933), 689—708].

R. Weitzenböck, Über die Matrixgleichung $X^2 = A$ [Proceedings Amsterdam **35** (1932), 328—329].

Ph. Franklin, Algebraic matric equations [Journal Math. Physics Massachusetts **10** (1932), 289—314].

D. E. Rutherford [Proc. Edinburgh Math. Soc. (2) **3** (1932), 135—143].

C. C. Mac Duffee, The theory of matrices [Ergebnisse d. Math. **2** (1933), Heft 5, 89—97 und die dort zitierte weitere Literatur].

Lösungen in $R(A)$ kennen, gehen wir für geeignete A zu allgemeineren Wurzeln über, die nicht mehr rational von A abhängen.

§ 1. *Die Zerlegung des Ringes* $R(A)$.

Bekanntlich genügt jede Matrix M einer eindeutig bestimmten Gleichung niedrigsten Grades:

$$\psi(M) = M^{\nu} + c_1 M^{\nu-1} + \ldots + c_{\nu} = 0.$$

Die Gleichung $\psi(\lambda) = 0$ heißt die zur Matrix M gehörige reduzierte charakteristische Gleichung oder auch die zu M gehörige Minimalgleichung, das Polynom $\psi(\lambda)$ die reduzierte charakteristische Funktion von M. Es ist $\psi(\lambda)$ der letzte Elementarteiler $E_k(\lambda)$ der Matrix $\lambda E - M$, und außerdem ist $\psi(\lambda)$ ein solcher Teiler der charakteristischen Funktion $\chi(\lambda) = |\lambda E - M| =$ Det $(\lambda E - M)$, der jeden Linearfaktor von $\chi(\lambda)$ wenigstens einmal enthält. Ist nun $g(\lambda)$ eine beliebige ganze rationale Funktion von λ, dann besitzt die Matrix $g(M)$ dann und nur dann eine verschwindende Determinante, wenn $g(\lambda)$ und $\psi(\lambda)$ nicht teilerfremd sind. Der Bereich $R(M)$ der rationalen Funktionen $\frac{g(M)}{h(M)}$ von M, wo die zur Matrix $h(M)$ gehörende Funktion $h(\lambda)$ zur reduzierten charakteristischen Funktion teilerfremd ist, ist isomorph zum Restklassenring $K[\lambda; \bmod \psi(\lambda)]$, wenn man nur diejenigen Brüche $\frac{g(\lambda)}{h(\lambda)}$ nach dem Modul $\psi(\lambda)$ zuläßt, deren Nenner in der reduzierten Form zum Modul teilerfremd ist; dem Element c_i des Koeffizientenkörpers K ist vermöge der Isomorphie das Skalarsystem

$$\begin{pmatrix} c_i & 0 \ldots 0 \\ 0 & c_i \\ \vdots & \ddots \\ 0 & \ddots\, c_i \end{pmatrix}$$

zugeordnet [2]). Die Untersuchung von Polynomen auf Nullstellen im Ringe $R(M)$ werden wir demgemäß dadurch führen, daß wir zum isomorphen Bereich der rationalen Funktionen modulo $\psi(\lambda)$ übergehen, da die Behandlung von $R(M)$ dadurch wesentlich vereinfacht wird.

Zu einer weiteren Vereinfachung gelangen wir dadurch, daß wir den Restklassenring mit dem Modul $\psi(\lambda) = \prod_{i=1}^{\mu} (\lambda - \alpha_i)^{k_i}$ additiv zerlegen in eine direkte Summe von Ringen

$$\overline{\mathfrak{R}}_i \backsimeq K[\lambda_i; \bmod (\lambda - \alpha_i)^{k_i}] \text{ [3]}).$$

[2]) Vgl. etwa: O. Haupt, Einführung in die Algebra II (Leipzig 1929), 628.

[3]) Vgl. c.d. v. d. Waerden, Moderne Algebra II, 47.

Ist $\psi(\lambda) = (\lambda - \alpha_i)^{k_i} \cdot \pi_i(\lambda)$, so sind die μ Funktionen $\pi_i(\lambda)$ zu einander teilerfremd. Es lassen sich nach dem Euklidschen Verfahren μ Multiplikatoren $\varrho_i(\lambda)$ so bestimmen, daß $\sum_{i=1}^{\mu} \varrho_i(\lambda)\pi_i(\lambda) = 1$ ist [4]).

Durch die μ Einheitsfunktionen $e_i(\lambda) = \varrho_i(\lambda)\pi_i(\lambda)$ läßt sich nun jede Funktion $\xi(\lambda)$ eindeutig in der additiven Form $\xi(\lambda) = \sum \xi_i(\lambda) e_i(\lambda)$ darstellen.

Es ist $e_i(\lambda)$ das Einselement im Ringe $K[\lambda; \bmod (\lambda - \alpha_i)^{k_i}]$ und für die $e_i(\lambda)$ gelten die Relationen:

$$\text{(a)} \qquad \sum_{i=1}^{\mu} e_i(\lambda) \equiv 1, \quad e_i^2(\lambda) \equiv e_i(\lambda) \bmod \psi(\lambda) \qquad \text{und}$$

$$e_\mu(\lambda) e_\nu(\lambda) \equiv 0 \bmod \psi(\lambda) \text{ für } \nu \neq \mu.$$

Es ist

$$e_i(\lambda) \equiv 1 \bmod (\lambda - \alpha_i)^{k_i},$$

weil $e_j(\lambda)$ durch $\pi_j(\lambda)$ teilbar, also

$$\equiv 0 \bmod (\lambda - \alpha_i)^{k_i}, \text{ wenn } j \neq i;$$

$$e_i^2(\lambda) \equiv e_i(\lambda),$$

denn die Differenz

$$e_i^2(\lambda) - e_i(\lambda) = e_i(\lambda)\big(e_i(\lambda) - 1\big)$$

ist durch $\psi(\lambda)$ teilbar, weil $e_i(\lambda)$ durch $\pi_i(\lambda)$ und wegen

$$e_i(\lambda) - 1 \equiv 1 - 1 = 0 \bmod (\lambda - \alpha_i)^{k_i}$$

auch durch $(\lambda - \alpha_i)^{k_i}$ teilbar ist. Daß $e_\mu(\lambda) e_\nu(\lambda) \equiv 0 \bmod \psi(\lambda)$ für $\nu \neq \mu$, ist klar. *Damit ist das allgemeine Problem der Betrachtung des Ringes $R(A)$ reduziert auf die Untersuchung eines Kongruenzrings, dessen Modul die Potenz eines Linearfaktors der Minimalgleichung ist.*

§ 2. *Die Untersuchung von $P(X, A)$ in einem Teilring. Hinreichende Bedingung für die Lösbarkeit von $P(X, A) = 0$.*

Jeder Lösung $\xi(A)$ von $P(X, A) = 0$ entspricht im Restklassenring eine Funktion $\xi(\lambda)$, welche die Kongruenz

$$\text{(1a)} \qquad P(x; \lambda) = x^l + a_1 x^{l-1} + \ldots + a_l \equiv 0 \bmod \psi(\lambda)$$

erfüllt und umgekehrt.

---

[4]) Wegen der „Kovarianten von Frobenius" $\varrho_i(\lambda)\pi_i(\lambda)$ vgl. FROBENIUS, Über vertauschbare Matrizen [Sitzungsberichte Akad. Berlin **1896**, 601] und auch J. WELLSTEIN [Arch. d. Math. u. Physik **5** (1903), 233].

Nun genügt aber jede Lösung $x(\lambda)$ von (1a) a fortiori der Kongruenz

(1b) $$P(x;\lambda) \equiv 0 \bmod (\lambda-\alpha_i)^{k_i}, \quad i=1,2,\ldots,\mu.$$

Bilden $x_1(\lambda), x_2(\lambda), \ldots, x_\mu(\lambda)$ ein Lösungssystem der μ Kongruenzen (1b), so ist infolge von (a) auch $x(\lambda)=\sum x_j(\lambda)e_j(\lambda) \bmod \psi(\lambda)$ eine Lösung von (1a). Demnach erhalten wir die Lösungen der Matrixgleichung in der Form

$$x(A)=\sum_j x_j(A)e_j(A).$$

Die Untersuchung der Gleichung $P(A;X)=0$ auf Lösungen im Ring der rationalen Funktionen der Matrix A ist demnach auf eine Untersuchung im Komponentenring zurückgeführt.

Wir betrachten die Fälle, daß die Minimalgleichung einfache oder mehrfache Wurzeln hat, getrennt.

A) Das reduzierte charakteristische Polynom hat lauter einfache Nullstellen. Da zur Lösbarkeit von $P(x;\lambda) \equiv 0 \bmod \psi(\lambda)$ die Lösbarkeit von $P(x;\lambda)\equiv 0$ im Ring $\overline{\mathfrak{R}}_i$ genügt, kann man im Falle $k=1$ sofort erkennen, daß (1) immer Nullstellen hat. Denn für $k=1$ ist jede Wurzel der Gleichung $P(x_i;\alpha_i)=0$ eine Lösungsfunktion von $P(x;\lambda)\equiv 0 \bmod \lambda-\alpha_i$, und jede Lösung muß offenbar einer dieser Wurzeln mod $(\lambda-\alpha_i)$ kongruent sein.

Satz 1: *Hat das reduzierte charakteristische Polynom der Matrix A lauter einfache Nullstellen, dann hat die Matrixgleichung $P(X,A)=0$ immer Lösungen.*

B) Das reduzierte charakteristische Polynom hat mehrfache Nullstellen. Wir können uns auf die Betrachtung des Polynoms (1) in einem Komponentenring beschränken. Wir wollen zunächst einen Satz beweisen, der zwar keine notwendige und hinreichende Bedingung für die Lösbarkeit von $P(X,A)=0$ ist, der aber doch die Frage mit einer sehr seltenen Ausnahme erledigt:

Hilfssatz: *Die Kongruenz $P(x;\lambda)\equiv 0$ mod $(\lambda-\alpha)^k$ ist für $k>1$ lösbar, wenn nicht $P(x;\alpha)$ Potenz einer anderen ganzen Funktion von x ist. Die Anzahl der α, für die bei gegebenen $P(x;\lambda)$ die Lösbarkeit fraglich bleibt, ist endlich. Diese α müssen gewisse der Wurzeln der aus den $a_i(\lambda)$ gebildeten Diskriminante sein.*

Zum Beweise des Satzes wollen wir untersuchen, wann $P(x;\lambda)\equiv 0 \bmod (\lambda-\alpha)^k$ für jedes noch so große k lösbar ist. Zu diesem Zwecke führen wir den Körper $K(\mathfrak{P}(\lambda-\alpha))$ aller Potenzreihen von $(\lambda-\alpha)$, die nur eine endliche Anzahl Glieder negativer Ordnung besitzen, ein. Die Koeffizienten in diesen Potenzreihen sollen Zahlen sein. Die Gesamtheit dieser Potenzreihen $\mathfrak{P}(\lambda-\alpha)$

bildet einen Körper. Diese Einführung ist genau analog zur Definition der p-adischen Zahlen zu verstehen [5]). Die Begriffe wie Näherungswerte, Kongruenz und Gleichheit werden wörtlich übertragen; Konvergenz und Divergenz dieser Reihen ist demgemäß gar nicht zu definieren. Wir fragen nun, wann die Gleichung $P(x; \lambda) = 0$ in $K(\mathfrak{P}(\lambda - \alpha))$ ganze Wurzeln hat. Sind solche Wurzeln vorhanden, so ist die Kongruenz eben für jedes noch so große k lösbar, und der $(k-1)^{\text{te}}$ Näherungswert der Wurzeln ist eine Lösung der Kongruenz. Kennen wir geeignete Näherungswerte dieser Wurzeln, dann ist es oft möglich, aus ihnen nach dem Newtonschen Verfahren die Wurzeln selbst zu finden. Können wir ferner von einem Anfangswert aus das Näherungsverfahren beliebig oft anwenden, dann ist die Existenz einer Wurzel von $P(x; \lambda) = 0$ in $K(\mathfrak{P}(\lambda - \alpha))$ gesichert.

Die nullten Näherungswerte der eventuell vorhandenen Wurzeln sind die Lösungen von $P(x; \lambda) \equiv 0 \bmod \lambda - \alpha$. Das sind aber die Wurzeln von $P(x; \alpha) = 0$ der Größe nach. Es sei nun η_0 eine von ihnen, also $P(\eta_0; \alpha) = 0$. Wir wollen nun zeigen, daß diesem η_0 fast immer eine Wurzel von $P(x; \lambda) = 0$ in $K(\mathfrak{P}(\lambda - \alpha))$ entspricht. Es ist

$$P(\eta_0 + h; \lambda) = P(\eta_0; \lambda) + hP'(\eta_0; \lambda) + \ldots + \frac{h^s}{s!} P^{(s)}(\eta_0; \lambda),$$

die Ableitungen sind dabei partiell nach x genommen.

Ist nun $P^{(i)}(\eta_0; \lambda)$ genau teilbar durch $(\lambda - \alpha)^{\varrho^{(i)}}$, $i = 0, 1, \ldots, s$, dann liefert der Ansatz

$$h = -\frac{P(\eta_0; \lambda)}{P'(\eta_0; \lambda)}$$

für den Bereich $K(\mathfrak{P}(\lambda - \alpha))$ einen besseren Wert, wenn $\varrho > \frac{i\varrho' - \varrho^{(i)}}{i-1}$ für $i = 2, 3, \ldots, s$ ist. Ist nun $\varrho' = 0$, so ist diese Bedingung erfüllt, da $\varrho \geqq 1$ und die rechte Seite der Ungleichung dann $\leqq 0$ ist. Die Ordnung von h selbst ist dann ϱ, also > 0. Nun ist aber die Ordnung von $P'(\eta_0 + h; \lambda)$ offenbar wieder Null, da ja $P'(\eta_0 + h; \lambda) \equiv P'(\eta_0; \lambda) \bmod (\lambda - \alpha)$ ist. Daher kann das Näherungsverfahren beliebig oft angewandt werden. Es existiert also in $K(\mathfrak{P}(\lambda - \alpha))$ eine ganze Wurzel von $\mathfrak{P}(x; \lambda) = 0$, wenn ϱ' Null ist.

---

[5]) Vgl. K. Hensel, Theorie der algebraischen Zahlen [Leipzig 1909], bezügl. des Beweises vgl. 72—73.

Nun war aber η_0 eine beliebige Wurzel von $P(x; \lambda) = 0$. Dieser Existenzbeweis für Wurzeln in $K(\mathfrak{P}(\lambda - \alpha))$ versagt also nur dann, wenn $P(x, \alpha) = 0$ *nur* mehrfache Wurzeln hat, da ja für eine einfache Wurzel η_0 immer $P'(\eta_0; \alpha) \neq 0$ ist. Es existiert also in $K(\mathfrak{P}(\lambda - \alpha))$ immer eine Wurzel von $P(x; \lambda) = 0$, wenn nicht $P(x; \alpha)$ Potenz einer anderen ganzen Funktion von x ist. Nach einem bekannten Satz [6]) ist das aber rational entscheidbar. Auch in diesem Ausnahmefall kann man bei einem bestimmt vorgegebenen $P(x; \lambda)$ noch immer versuchen, ob nicht das Näherungsverfahren doch zum Ziel führt. Für jede gegebene Funktion ist die Anzahl dieser Ausnahmefälle für α immer endlich, da sie ja alle unter den Wurzeln der Diskriminante von $P(x; \lambda)$ vorkommen müssen. Damit ist der behauptete Satz bewiesen. In Bezug auf den Ring $R(A)$ können wir daher den Satz aussprechen:

Satz 2: *Die Matrixgleichung* $P(X, A) = 0$ *hat immer eine Lösung, wenn nicht* $P(x, \alpha)$ *Potenz einer anderen ganzen Funktion von* x *ist, wobei* α *eine Wurzel der Minimalgleichung für die Matrix* A *ist.*

§ 3. *Spezielle Matrixgleichungen.*

A) *Die Gleichung* $X^2 + g(A)X + h(A) = 0$.

Wegen der vielseitigen Behandlung, welche die Matrixgleichung $X^2 = A$, wo A eine beliebige quadratische Matrix $A = (a_{ik})$ bedeutet, in der Literatur [7]) erfahren hat, beginnen wir zunächst mit der Matrixgleichung $X^2 + g(A)X + h(A) = 0$, wobei $g(A)$ bzw. $h(A)$ diejenige Matrix bedeutet, die dadurch entsteht, daß man in den Polynomen $g(\lambda) = \sum_{i=0}^{\mu} \alpha_i \lambda^i$ bzw. $h(\lambda) = \sum_{i=1}^{\mu} \beta_i \lambda^i$ an Stelle von λ die Matrix A und von α_i bzw. β_i skalare Matrizen setzt.

---

[6]) Vgl. E. Steinitz: Algebraische Theorie der Körper [Journal f. r. u. angew. Math. **137** (1910), 167—309 (§ 10, Satz 5)] neu herausgegeben [Berlin 1930].

[7]) Cayley, Memoir on the theory of matrices [Phil. Trans. Royal Soc. London **148** (1858)] gibt eine Lösung der Gl. $X^2 = A$, wo A eine bekannte Matrix zweiter oder dritter Ordnung ist.

Sylvester, Sur les puissances et les racines de substitutions bilinéaires [C. R. **94** (1882)].

G. Frobenius [Sitzungsberichte Akad. Berlin **1896**, 7—16] u. auch M. Bôcher, Einführung in die höhere Algebra [Teubner 1910], 322, derselbe Beweis verallgemeinert auf $X^n = A$, A nicht singulär, findet sich bei Dickson-Bodewig, Höhere Algebra [Teubner 1929], 107—109.

R. Weitzenböck, l. c. [1]). Gibt allgemeinere Lösungen. Methode ist anwendbar auf $X^n = A$, $n > 2$.

Die Minimalgleichung für A sei von der Gestalt $(\lambda-\alpha)^k$. Es handelt sich also um die Kongruenz

$$x^2+g(\lambda)x+h(\lambda) \equiv 0 \bmod (\lambda-\alpha)^k. \tag{a}$$

Eine etwa vorhandene Lösung ist eine Restfunktion der Form $\xi(\lambda)=\xi_0+\xi_1(\lambda)+\ldots+\xi_{k-1}\lambda^{k-1}$. Um die Koeffizienten ξ_i zu bestimmen, setzen wir in der Kongruenz (a) für λ den Wert α ein und erhalten als 1. Gleichung

$$\begin{aligned}(\xi_0+\xi_1\alpha+\ldots+\xi_{k-1}\alpha^{k-1})^2+\\ +(a_0+\alpha_1\alpha+\ldots+a_{k-1}\alpha^{k-1})(\xi_0+\ldots+\xi_{k-1}\alpha^{k-1})+\\ +\beta_0+\ldots+\beta_{k-1}\alpha^{k-1}=0.\end{aligned}$$

Ist $k=1$, dann ist immer eine Lösung vorhanden. Es sei $k>1$. Damit $x^2(\lambda)+g(\lambda)x(\lambda)+h(\lambda)$ durch $(\lambda-\alpha)^k$ teilbar ist, muß $2x(\lambda)x'(\lambda)+g(\lambda)x'(\lambda)+g'(\lambda)x(\lambda)+h'(\lambda)$ durch $(\lambda-\alpha)^{k-1}$ teilbar, also $2x(\alpha)x'(\alpha)+\ldots+h'(\alpha)=0$ sein. Daraus ergibt sich zur Bestimmung der ξ_i

$$x'(\alpha)=\frac{-g'(\alpha)x(\alpha)-h'(\alpha)}{2x(\alpha)+g(\alpha)}$$

und wegen

$$x(\alpha)=-\frac{g(\alpha)}{2}\pm\sqrt{\frac{g^2(\alpha)}{4}-h(\alpha)}$$

$$\xi_1+2\xi_2\alpha+\ldots+(k-1)\xi_{k-1}\alpha^{k-1}=x'(\alpha)=\frac{-g'(\alpha)x(\alpha)-h'(\alpha)}{\pm\sqrt{g^2(\alpha)-4h(\alpha)}}=$$

$$=\frac{-\frac{gg'}{2}\pm\frac{g'}{2}\sqrt{\ }}{\pm\sqrt{g^2-4h}}$$ als 2. Gleichung. Gehen wir bis zur $(k-1)^{\text{ten}}$ Ableitung, dann haben wir die erforderlichen k Gleichungen zur Bestimmung der ξ_i. Wir sehen nun sofort, daß die Lösbarkeit der Kongruenz (a) nur dann in Frage gestellt ist, wenn die aus den Koeffizienten gebildete Diskriminante $g^2(\lambda)-4h(\lambda)=\Delta(\lambda)$ für die Nullstelle $\lambda=\alpha$ des Moduls verschwindet. Ist $\Delta(\alpha)\neq 0$, dann ist die Kongruenz immer lösbar. Wir kennen somit die Bedingungen der Lösbarkeit der quadratischen Gleichung im Ringe $\overline{\mathfrak{R}}_i$, also auch im Ringe $K[\lambda;\ \mathrm{mod}\ \psi(\lambda)]$ und formulieren

Satz 3: Die Matrixgleichung $X^2+g(A)X+h(A)=0$ ist immer lösbar, wenn die Minimalgleichung für A nur einfache Nullstellen hat, oder wenn die Funktion $g^2(\lambda)-4h(\lambda)$ für die mehrfachen Nullstellen der Gleichung niedrigsten Grades für A nicht verschwindet.

Ist nun speziell $g(A)$ eine Skalarmatrix und bezeichnet man $h(A)$ mit M, dann ist das zur Matrixgleichung $X^2+cX+M=0$

gehörige Gleichungssystem, wenn $\mu_1, \ldots, \mu_\varrho$ die Nullstellen des Minimalpolynoms für M bedeuten, von der Form

$$\xi_1^{(i)}+\xi_1^{(i)}\mu_i+\ldots+\xi_{k_i-1}^{(i)}\cdot\mu_i^{k_i-1}=\frac{-c+\sqrt{c^2-4\mu_i}}{2}$$

$$\xi_1^{(i)}+2\xi_2^{(i)}\mu_i\ldots=\frac{-1}{\sqrt{c^2-4\mu_i}}$$

usw. bis zur $(k_i-1)^{\text{ten}}$ Ableitung, $i=1, 2, \ldots, \varrho$.

Die Determinante des Gleichungssystems ist $\neq 0$, und das Gleichungssystem ist dann und nur dann lösbar, wenn $c^2-4\mu_i\neq 0$ ist.

(So hat z.B. $X^2+2X+A=0$ im Ring der rationalen Funktionen von

$$A=\begin{pmatrix}1&0&0\\1&1&0\\0&1&1\end{pmatrix}$$

keine Lösung, da die Gleichung niedrigsten Grades für A lautet $(1-\lambda)^2=0$ und $\Delta(1)=0$ ist, während $X^2-A=0$ durch

$$X=\begin{pmatrix}1&0&0\\\frac{1}{2}&1&0\\-\frac{1}{8}&\frac{1}{2}&1\end{pmatrix}$$

erfüllt wird.)

Ist $c=0$ und $h(A)=-A$, dann liegt die Matrixgleichung $X^2=A$ vor und es ergibt sich

Satz 4: *Die Matrixgleichung $X^2=A$ ist im Ringe der rationalen Funktionen von A dann und nur dann lösbar, wenn die Gleichung niedrigsten Grades $\psi(\lambda)$ für A höchstens durch die erste Potenz von λ teilbar ist.*

Wir beweisen Satz 4 unter Zuhilfenahme von Satz 2 und zeigen außerdem:

Satz 5: *Hat das Minimalpolynom für A die in* § 1 *angegebene Zerlegung, und haben die $e_i(\lambda)$ die dort festgesetzte Bedeutung, dann lassen sich die Lösungen der Matrixgleichung $X^2=A$ explizit in der Form angeben:*

$$x(A)=\sum_{i=1}^{\nu}e_i(A)\sum_{l=0}^{k_i-1}\binom{\frac{1}{2}}{l}\alpha_i^{\frac{1}{2}-l}\cdot E(A-\alpha_i E)^l.$$

Beweis: Aus § 2B ergibt sich für $P(x;\lambda)=x^2-\lambda\equiv 0 \bmod (\lambda-\alpha)^k$, daß $P(x;\alpha)=x^2-\alpha$ ist. Das ist aber offenbar dann und nur dann die Potenz einer Funktion von x, wenn $\alpha=0$ ist. Nur dann ist also die Lösbarkeit fraglich. Man sieht, daß schon $x^2-\lambda\equiv 0 \bmod \lambda^2$ nicht lösbar ist, da ja x durch λ teilbar sein müßte und daraus nach Division durch λ folgen würde $1\equiv 0 \bmod \lambda$. In allen anderen Fällen existiert also eine Lösung in $K[\mathfrak{P}(\lambda-\alpha)]$, und

der Ansatz $(\xi_0+\xi_1(\lambda-\alpha)+\ldots)^2=\lambda$ ergibt mit Hilfe der Substitution $\lambda=\alpha+\alpha\mu$ sofort $\xi_0+\xi_1\alpha\mu+\ldots=\alpha^{\frac{1}{2}}\cdot(1+\mu)^{\frac{1}{2}}$. Da die Wurzel rechts in $K(\mu)$ enthalten ist, ist die Binomialentwicklung erlaubt und durch Koeffizientenvergleichung erhält man $\xi_i=\binom{\frac{1}{2}}{i}\alpha^{\frac{1}{2}-i}$. Daher erhält man den obigen Ansatz für x, der auch für $\alpha_1=0$ und $k_1=1$ sich als richtig erweist.

Wir kommen auf die Gleichung $X^2=A$ in 4B wieder zurück.

B) *Die Gleichung* $X^n=A^m$. Um diese Gleichungen zu untersuchen, befassen wir uns mit binomischen Kongruenzen. Denken wir an die Analogie dieser Aufgabe mit der Theorie der Potenzreste in der Zahlentheorie, so stellen wir fest, daß im Restklassenring $K[\lambda;\ \mathrm{mod}\ \psi(\lambda)]$ die Theorie der Potenzreste sich ganz einfach gestaltet im Gegensatz zum zahlentheoretischen Analogen. Zur Untersuchung der binomischen Kongruenzen $x^n-a(\lambda)\equiv 0 \mod \psi(\lambda)$ greifen wir wieder auf Satz 2 zurück. $x^n-a(\alpha)$ ist sicher keine Potenz einer anderen ganzen Funktion von x, wenn $a(\lambda)$ für die Nullstelle α des Moduls nicht verschwindet. Dementsprechend setzen wir $a(\lambda)$ und $\psi(\lambda)$ als teilerfremd voraus.

Unter dieser Voraussetzung läßt sich die Kongruenz einfach erledigen: Die gesuchte Lösung $x(\lambda)$ sei

$$\equiv \xi(\lambda)=\xi_0+\xi_1\lambda+\ldots+\xi_{k-1}\lambda^{k-1} \mod (\lambda-\alpha)^k.$$

Aus $(\xi(\lambda))^n-a(\lambda)\equiv 0 \mod (\lambda-\alpha)^k$ folgt zunächst für $\lambda=\alpha$ die Gleichung: $\xi(\alpha)=\overset{\frac{1}{n}}{a(\alpha)}$.

Da α eine k-fache Nullstelle des Moduls ist, also die σ^{te} Ableitung von $\xi(\lambda)^n-a(\lambda)$ durch $(\lambda-\alpha)^{k-\sigma}$ teilbar ist, erhält man zur Bestimmung der ξ_i das folgende Gleichungssystem

$$(\xi(\alpha))^{(\varrho)}=\left(\overset{\frac{1}{n}}{a(\alpha)}\right)^{(\varrho)},\quad \varrho=0,\ \ldots,\ k-1,$$

wobei die Differentiation sich auf α bezieht und der Wert $\left(\overset{\frac{1}{n}}{a(\alpha)}\right)^{(\varrho)}$ in der Weise zu bilden ist, daß man in die ϱ^{te} Ableitung der Funktion $\left(a(\lambda)\right)^{\frac{1}{n}}$ nach λ für λ den Wert α einsetzt.

Die Determinante des Gleichungssystems ist

$$D=\begin{vmatrix} 1 & \alpha & \alpha^2 & \ldots & \alpha^{k-1} \\ 0 & 1 & 2\alpha & \ldots & (k-1)\alpha^{k-2} \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & \ldots & \ldots & \ldots & (k-1)! \end{vmatrix},$$

also sicher $\neq 0$, wie auch α gewählt sein mag.

Es sei $a(\lambda)=(\lambda-\beta)^m$, wo $m<k$ und positiv ganz. Dann gilt:

Hilfssatz: *Die Kongruenz* $x^n-(\lambda-\beta)^m\equiv 0 \bmod (\lambda-\alpha)^k$ *ist für* $\beta\neq\alpha$ *immer lösbar, für* $\beta=\alpha$ *dann und nur dann, wenn* n/m *ist.*

Beweis: Für $a(\lambda)=(\lambda-\beta)^m$ erhält man als Ableitungsfolge:

$$\left(a(\alpha)^{\frac{1}{n}}\right)^{(\varrho)}=\left(\frac{m}{n}-\varrho+1\right)\ldots\frac{m}{n}(\lambda-\beta)^{\frac{m}{n}-\varrho},\quad \varrho=1,\ \ldots,\ k-1.$$

Die Lösbarkeit ist nur fraglich für $\alpha=\beta$.

a) Sei n/m, also $\frac{m}{n}$ ganz. Dann ist die $\frac{m}{n}$-te Ableitung für $\lambda=\alpha$ von Null verschieden und hat den Wert $\frac{m}{n}!$ Alle ξ_i sind eindeutig bestimmt. Es ist immer eine Lösung vorhanden.

b) $\frac{m}{n}<1$; dann ist $\frac{m}{n}-\varrho<0$, also $\left(\frac{1}{\alpha^n(\lambda)}\right)^{(\varrho)}$ für $\lambda=\alpha$ sinnlos.

c) $\frac{m}{n}>1$; dann existiert $\left(a(\lambda)^{\frac{1}{n}}\right)^{k-1}$, wenn $\frac{m}{n}-k+1\geqq 0$ ist; wegen $m<k$ erhält man $n<\frac{k}{k-1}$. Da $k>1$ vorausgesetzt ist, folgt $n=1$.

Wählen wir $a(\lambda)$ nun beliebig in $K[\lambda;\ \mathrm{mod}\ \psi(\lambda)]$, dann ist für $(a(\lambda);\ (\lambda-\alpha)^k)=1$ immer eine Lösung vorhanden. Hat $a(\lambda)$ mit dem Modul einen gemeinsamen Teiler, ist also $a(\alpha)=0$, dann hat die Kongruenz $x^n-a(\lambda)\equiv 0 \bmod (\lambda-\alpha)^k$ dann und nur dann eine Lösung, wenn $a(\lambda)=(\lambda-\alpha)^m\cdot g(\lambda)$; $g(\alpha)\neq 0$ und n/m ist.

Satz 6: *Die Matrizengleichung* $X^n=A^m$ *ist immer lösbar, wenn* A *nicht-singulär ist. Bei singulärem* A, *wobei die Gleichung niedrigsten Grades für* A *die* k*-fache Wurzel* $\alpha=0$ *hat, sind auch immer Lösungen vorhanden, wenn* $k\leqq m$ *ist, für* $k>m$ *dagegen dann und nur dann, wenn* n *ein Teiler von* m *ist.*

C) *Die Gleichung* $P(X)=A$. Der Übergang zur allgemeineren Matrixgleichung $P(X)=A$, wobei $P(x)$ ein normiertes Polynom mit Skalarkoeffizienten, dem Absolutglied Null und A eine beliebige Matrix ist, bietet nach Satz 2 keine Schwierigkeit. Wir stellen ohne Weiteres die folgende hinreichende Bedingung für die Lösbarkeit fest:

Sei $\psi(\lambda)$ ein Polynom vom Grade $\nu>1$ in λ mit den verschiedenen Nullstellen $\alpha_1, \alpha_2, \ldots, \alpha_\mu$, $P(x)$ sei ein normiertes Polynom vom Grade $n>1$ in x, dessen absolutes Glied Null ist. Die Gleichung $P(x)-\alpha_j=0$ $(j=1,\ \ldots,\ \mu)$ habe wenigstens eine einfache Wurzel für jedes α_j, das eine mehrfache Wurzel von

$\psi(\lambda)=0$ ist, und es sei $\psi(A)=0$ die Minimalgleichung für die quadratische Matrix A der Ordnung m, dann existiert sicher eine Matrix X der Ordnung m, so daß $P(X)=A$ ist, wobei X als Polynom in A mit skalaren Koeffizienten ausdrückbar ist.

Mit Hilfe von einfachen Kongruenzbetrachtungen überzeugt man sich leicht, daß die im vorhergehenden Satz ausgesprochene hinreichende Bedingung für die Gleichung $P(X)=A$ auch zur Lösbarkeit notwendig ist. Denn es enthalte $P(x)-\alpha=0$ nur mehrfache Wurzeln. Wir entnehmen dem Körper $K(\mathfrak{P}(\lambda-\alpha))$ das Element $\sum_{i=0}^{k_i-1}\xi_i(\lambda-\alpha)^i$. Es soll der Kongruenz (1b) genügen. Für $\lambda=\alpha$ folgt, daß $\xi_0^t+a_1\xi_0^{t-1}+\ldots+a_{t-1}\xi_0-\alpha=0$ sein muß. Da die Ableitung notwendig durch $(\lambda-\alpha)^{k-1}$ teilbar sein muß, folgt aus (1b) und dem Vorhandensein von lauter mehrfachen Wurzeln von $P(x)-\alpha_j=0$ die Gleichung $-1=0$. Nachdem wir so die notwendigen und hinreichenden Bedingungen für die Lösbarkeit von $P(X)=A$ kennen, ergibt sich aus der Darstellung $X(A)=\sum_{i=1}^{\mu}x_i(A)e_i(A)$ sofort die Anzahl der Lösungen. Ist nämlich σ_j die Anzahl der verschiedenen Wurzeln von $P(x)-\alpha_j=0$, wenn α_j einfache Wurzel der Minimalgleichung bzw. die Zahl der einfachen Wurzeln von $P(x)-\alpha_j=0$, wenn α_j mehrfache Wurzel der reduzierten charakteristischen Gleichung ist, dann gibt es — je nach der Wahl der $x_i(A)$ — im Ganzen $\prod_{j=1}^{\mu}\sigma_j$ Lösungen. Zusammenfassend haben wir für die Gleichung $P(X)=A$ den

Satz 7: Für die Lösbarkeit der Matrixgleichung $P(X)=A$ ist notwendig und hinreichend, daß das Polynom $P(x)-\lambda$ für die mehrfachen Wurzeln der Minimalgleichung wenigstens eine einfache Wurzel hat. Ist σ_j die Anzahl dieser einfachen Wurzeln bzw. die Anzahl der verschiedenen Wurzeln von $P(x)-\lambda=0$ für die einfachen Wurzeln der Minimalgleichung, dann ist die Anzahl der Lösungen gleich $\prod_{j=1}^{\mu}\sigma_j$.

D) *Skalargleichungen. Zerlegung von Polynomen.* Wir beschränken uns nun auf solche Polynome, deren Koeffizienten lauter Skalarmatrizen n^{ter} Ordnung sind: $X^t+s_1X^{t-1}+\ldots+s_t=0$.

In dem zu einer Matrix A der n^{ten} Ordnung gehörigen Ringe $R(A)$ besitzt eine derartige Gleichung zunächst t Skalarmatrizen als „Gaußsche" Wurzeln. Hat die Minimalgleichung lauter verschiedene Wurzeln α_ϱ, dann hat man wegen $k_i=1$ in jedem der ν Teilringe einfach die t Nullstellen des Polynoms $x^t+s_1x^{t-1}+\ldots+s_t$,

und es ist $x(A)=\sum_{i=1}^{\nu}\beta_j e_i(A)$. Da die Koeffizienten β_j der Einheiten $e_i(A)$ unabhängig voneinander t Werte durchlaufen, erhalten wir im Ganzen t^ν Lösungen. Sind in der Minimalgleichung $k_i > 1$ vorhanden und ist $\sum_{j=1}^{\mu} k_j = \nu$, dann erhält man entweder t^μ oder beliebig viele Lösungen, je nachdem ob die Gaußschen Nullstellen einfach oder mehrfach sind. Denn es seien:

a) alle β verschieden. Ist in einem Faktor $x(\lambda)-\beta_\tau$ ein Faktor $\lambda-\alpha_i$ enthalten, dann auch $(\lambda-\alpha_i)^{k_i}$, weil aus $x(\lambda)-\beta_{\tau_1}\equiv 0$, $x(\lambda)-\beta_{\tau_2}\equiv 0 \bmod \lambda-\alpha_i$ folgt $\beta_{\tau_1}=\beta_{\tau_2}$. Die Bestimmung der ξ_i in $x(\lambda)=\sum_{j=0}^{k_i-1}\xi_i\lambda^j$ ergibt $\xi_s^{(i)}=\begin{cases}\beta_j & \text{für } s=0\\ 0 & \text{,, } s>0.\end{cases}$ Demnach ist die Anzahl der $x(A)$ gleich t^μ.

b) nicht alle β sind verschieden. β_τ sei eine s-fache Wurzel. Dann ersieht man aus $(x(\lambda)-\beta_\tau)^s \equiv 0 \bmod (\lambda-\alpha_i)^{k_i}$, daß nicht mehr alle ξ_i eindeutig bestimmt sind. Daher:

Satz 8: *Ist μ die Anzahl der verschiedenen Wurzeln der Minimalgleichung und hat die Skalargleichung lauter verschiedene Gaußsche Wurzeln, dann ist die Gesamtzahl der Lösungen t^μ. Bei mehrfachen Wurzeln von Minimal- und Skalargleichung sind unendlich viele Lösungen vorhanden.*

Wählt man die Kombination, daß alle x_i einander gleich sind, nämlich alle gleich einer der t gewöhnlichen Gaußschen Wurzeln β der Gleichung t^{ten} Grades, dann hat man wegen $x(\lambda)=\beta\sum_{i=1}^{\mu}e_i(\lambda)$ und $\Sigma e_i(\lambda)=1$ für $x(\lambda)$ die t Gaußschen Nullstellen. Für eine andere Zusammenstellung der x_i erhält man weitere Lösungen, und zwar gehören diese nicht mehr dem Körper der komplexen Zahlen an. Je t der Lösungen aus dem Matrizenring kann man nun wiederum als „zueinander konjugiert" auffassen insofern, als die symmetrischen Grundfunktionen derselben die Eigenschaft haben, die Koeffizienten α_i der Gleichung t^{ten} Grades zu liefern.

Ist ν die Anzahl der voneinander verschiedenen Wurzeln der Gleichung niedrigsten Grades für A, so gibt es demnach im Ringe der rationalen Funktionen von A $t^{\nu-1}$ Zerlegungen eines Polynoms

$$S^t+a_1S^{t-1}+\ldots+a_t \tag{2}$$

in ein Produkt von t Linearfaktoren. Bezeichnen wir die Gaußschen Nullstellen von (2) der Reihe nach mit $s^{(0)}, s^{(1)}, \ldots, s^{(t-1)}$, und nennen wir $s^{(i)}$ die i^{te} Konjugierte zu $s^{(0)}$, dann entspricht der Lösung

$x^{(0)}(A) = s^{(i_1)}e_1(A) + \ldots + s^{(i_\mu)}e_\mu(A)$ als i^{te} Konjugierte
$x^{(i)}(A) = s^{(i_1+i)}e_1(A) + \ldots + s^{(i_\mu+i)}e_\mu(A)$,

wobei $i_\varkappa + i$ modulo t auf den kleinsten positiven Rest zu reduzieren ist. Denn

$$\sum_{j=0}^{t-1} x^{(j)}(A) = \Big(\sum_j s^{(j)}\Big)e_1(A) + \ldots + \Big(\sum_j s^{(j)}\Big)e_\mu(A) = \sum_j s^{(j)} \sum_j e_j(A) = -a_1$$

wegen $\sum_j s^{(j)} = -a_1$ und $\sum_j e_j(A) = 1$.

In analoger Weise beweist man die Richtigkeit der Behauptung unter Benutzung der Beziehungen $e_i^2(A) = e_i(A)$; $e_i(A) \cdot e_j(A) = 0$ $(j \neq i)$ und $\Sigma e_j(A) = 1$. Sei etwa $A = \begin{pmatrix} 2 & 0 \\ 0 & 0 \end{pmatrix}$, also $\psi(\lambda) = \lambda(\lambda - 2)$; $\alpha_1 = 2$; $\alpha_2 = 0$.

Wir zerlegen das Polynom $T^3 + 2T^2 - T - 2$. Die Anzahl der Nullstellen in $R(A)$ ist $t^\nu = 3^2 = 9$, nämlich

$$x_1 = 1; \quad x_2 = -1; \quad x_3 = -2;$$
$$x_4 = -1 + A = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}; \quad x_5 = -2 + \tfrac{1}{2}A = \begin{pmatrix} -1 & 0 \\ 0 & -2 \end{pmatrix}; \quad x_6 = 1 - \tfrac{3}{2}A = \begin{pmatrix} -2 & 0 \\ 0 & 1 \end{pmatrix};$$
$$x_7 = 1 - A = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}; \quad x_8 = -2 + \tfrac{3}{2}A = \begin{pmatrix} 1 & 0 \\ 0 & -2 \end{pmatrix}; \quad x_9 = -1 - \tfrac{1}{2}A = \begin{pmatrix} -2 & 0 \\ 0 & -1 \end{pmatrix};$$

und demgemäß:

$$\begin{aligned} T^3 + 2T^2 - T - 2 &= \Big[T - \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\Big]\Big[T - \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}\Big]\Big[T - \begin{pmatrix} -2 & 0 \\ 0 & -2 \end{pmatrix}\Big] \\ &= \Big[T - \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}\Big]\Big[T - \begin{pmatrix} -1 & 0 \\ 0 & -2 \end{pmatrix}\Big]\Big[T - \begin{pmatrix} -2 & 0 \\ 0 & 1 \end{pmatrix}\Big] \\ &= \Big[T - \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}\Big]\Big[T - \begin{pmatrix} -2 & 0 \\ 0 & -1 \end{pmatrix}\Big]\Big[T - \begin{pmatrix} 1 & 0 \\ 0 & -2 \end{pmatrix}\Big]. \end{aligned}$$

Satz 9: *Ist μ die Anzahl der verschiedenen Eigenwerte einer Matrix A, dann läßt sich jedes Polynom t^{ten} Grades mit einfachen Nullstellen im Ringe der rationalen Funktionen von A auf $t^{\mu-1}$ Arten in ein Produkt von t Linearfaktoren zerlegen.*

Satz 9a: *Von den t^μ Nullstellen eines Polynoms t^{ten} Grades im Ringe der rationalen Funktionen einer Matrix A sind je t zueinander konjugiert, d.h. die symmetrischen Grundfunktionen dieser t Nullstellen sind Skalarmatrizen.*

Bei der Zerlegung des Polynoms $P(T) - A = 0$ betrachten wir den Fall, daß $\psi(\lambda)$ lauter einfache Wurzeln hat. Es gibt dann nach dem Vorhergehenden t^ν Lösungen der Matrixgleichung $P(X) = A$ und entsprechend wieder $t^{\nu-1}$ Gruppen von je t konjugierten Matrizen. Bezeichnen wir die t Wurzeln von $P(x) - \alpha_j = 0$ mit $\xi_j^{(1)}, \ldots, \xi_j^{(t)}$ und werden die t zueinander konjugierten Lösungen genau so wie oben definiert, dann erkennt man sofort, daß die $t-1$ ersten symmetrischen Grundfunktionen Skalarmatrizen sind und zwar gerade die Skalarkoeffizienten von $P(x)$; die t^{te} Grundfunktion, nämlich das Produkt von t Konjugierten, ist

$\varphi(A) = \sum_{j=1}^{\nu} \left(\prod_{i=1}^{t} \xi_j^{(i)} \right) e_j(A)$. Es ist $\prod \xi_j^{(i)} = (-1)^{t+1}\alpha_j$, also $\varphi(A) = (-1)^{t+1} \cdot \Sigma \alpha_j e_j$. Die rechte Seite ist (abgesehen vom Vorzeichen) im Ringe $R(A)$ die Lösung der Gleichung $X = A$, d.h. $\varphi(A) = (-1)^{t+1}A$. Setzen wir voraus, daß $P(x) - \lambda = 0$ für $\lambda = \alpha_j$ lauter einfache Wurzeln hat bei mehrfachen Wurzeln der Minimalgleichung, dann gibt es t^{μ} Lösungen und entsprechend wieder $t^{\mu-1}$ Zerlegungen des Polynoms $P(X) - A$.

Satz 10: *Hat die Gleichung* $P(x) - \lambda = 0$ *für die Nullstellen des Minimalpolynoms lauter einfache Wurzeln, dann zerfällt das Polynom* $P(T) - A$ *auf* $t^{\mu-1}$ *Arten in ein Produkt von je t Linearfaktoren*, z.B. ist (vgl. die späteren Beispiele)

$$T^2 - 3T + \begin{pmatrix} 3 & -3 & -6 \\ \frac{7}{3} & -5 & -2 \\ \frac{4}{3} & -4 & 0 \end{pmatrix} = \left[T - \begin{pmatrix} \frac{9}{2} & -\frac{15}{2} & 3 \\ \frac{11}{6} & -\frac{7}{2} & 1 \\ \frac{4}{3} & -4 & 3 \end{pmatrix} \right] \left[T - \begin{pmatrix} -\frac{3}{2} & \frac{15}{2} & -3 \\ -\frac{11}{6} & \frac{13}{2} & -1 \\ -\frac{4}{3} & 4 & 0 \end{pmatrix} \right]$$

eine der vier möglichen Zerlegungen des Polynoms $T^2 - 3T + A = 0$, wo $A = \begin{pmatrix} -3 & 3 & 6 \\ -\frac{7}{3} & 5 & 2 \\ -\frac{4}{3} & 4 & 0 \end{pmatrix}$, in $R(A)$. Zu weiteren Zerlegungen (die natürlich dann außerhalb $R(A)$ führen) kann man bei passendem A durch die folgenden Betrachtungen geführt werden.

§ 4. *Weitere Lösungen von Matrixgleichungen.*

A. Bisher hatte es sich nur um die Bestimmung solcher Lösungen von Matrixgleichungen $P(X, A) = 0$ gehandelt, die selbst dem Ringe $R(A)$ angehörten. Zu weiteren Lösungen kann man durch folgende Überlegung kommen: Geht die Matrix $M = (m_{ik})$ durch kontragrediente Transformation in $N = (n_{ik})$ über, sind also M und N ähnlich, d.h. ist $SMS^{-1} = N$, so folgt leicht, daß $SM^iS^{-1} = N^i$ und $S(r_0 + r_1M + \ldots + r_{\nu-1}M^{\nu-1})S^{-1} = r_0 + r_1N + r_2N^2 + \ldots + r_{\nu-1}N^{\nu-1}$ also die rationalen Funktionen $g(M)$ und $g(N)$ ähnlich sind, denn aus $SMS^{-1} = N$ ergibt sich $SMS^{-1} \cdot SMS^{-1} = SM^2S^{-1} = N^2, \ldots$ und $\sum_j a_j SM^jS^{-1} = S(\sum_j a_j M^j)S^{-1} = \sum_j a_j N^j$. Aus $\psi(M) = 0$ folgt notwendig $\psi(N) = 0$ und umgekehrt, also genügen ähnliche Matrizen derselben Gleichung niedrigsten Grades.

Es sei die Matrix N von der Gestalt $N = \begin{pmatrix} N_1 & & & \\ & N_2 & & \\ & & \ddots & \\ & & & N_m \end{pmatrix}$, also bei N ist eine Reihe von „Feldern" längs der Diagonalen aneinandergereiht. (Eine bel. Matrix A z.B. ist stets einer Matrix der Form $\begin{pmatrix} A_1^* & & & \\ & A_2^* & & \\ & & \ddots & \\ & & & A_m^* \end{pmatrix} = B$ ähnlich, bei der A_i^* dadurch definiert ist, daß Det $(\lambda E - A_i^*) = 0$ die Gleichung niedrigsten Grades

darstellt, die durch die Matrix $\begin{pmatrix} A_i^* & & \\ & \ddots & \\ & & A_m^* \end{pmatrix}$ befriedigt wird.)

Kennt man Lösungen $Y_j^{(1)}, Y_j^{(2)}, \ldots$ der Gleichung $Y^t + r_1(N_j)Y^{t-1} + \ldots + r_t(N_j) = 0$, $(j = 1, 2, \ldots, m)$, $r_k(N_j)$ = rationale Funktion von N_j, dann erhält man in $\begin{pmatrix} Y_1^{(i_1)} & & \\ & \ddots & \\ & & Y_m^{(i_m)} \end{pmatrix}$ eine Lösung der Matrixgleichung $P(Y; N) = Y^t + r_1(N)Y^{t-1} + \ldots + r_t(N) = 0$. Durch Transformation gelangt man zu

$S^{-1}P(Y; N)S = X^t + r_1(M)X^{t-1} + \ldots + r_t(M) = 0$ mit $X = S^{-1}YS$.

Ist diese Gleichung in X vorgelegt, dann kann man M in N transformieren und die transformierte Gleichung behandeln. Kennt man die Lösungen Y der transformierten Gleichung, dann ist $X = S^{-1}YS$ sicher auch eine Lösung der gegebenen Gleichung. Auf diese Weise gelangt man zu allgemeineren Lösungen, eben solchen, die außerhalb des Ringes $R(M)$ liegen.

B) *Die Gleichung* $X^2 = A$. Für die Gleichung $X^2 = A$ wollen wir hieraus eine weitere Folgerung ziehen. Kennt man die in der oben angegebenen Art zu ermittelnden Lösungen der Gleichung $Y^2 = B$, so gewinnt man folgendermaßen weitere Werte für Y: Wir zerschneiden $\begin{pmatrix} Y_1 & & \\ & \ddots & \\ & & Y_m \end{pmatrix}$ auf alle möglichen Arten in vier Teilmatrizen derart, daß in der so entstehenden Matrix $\begin{pmatrix} Z_1 & O \\ O & Z_2 \end{pmatrix}$ Z_1 und Z_2 quadratische Matrizen sind. Wird dann U so gewählt, daß $Z_1U + UZ_2 = 0$ ist [8]), dann ist auch $Z = \begin{pmatrix} Z_1 & U \\ O & Z_2 \end{pmatrix}$ und entsprechend $\begin{pmatrix} Z_1 & O \\ V & Z_2 \end{pmatrix}$ eine Lösung von $Y^2 = B$ und $S^{-1}ZS = X$ eine Lösung von $X^2 = A$.

§ 5. *Die Zurückführung auf ein Eliminationsproblem und Beispiele.*

A) Im Anschlusse an das Vorhergehende untersuchen wir die Gleichung:

$$X^2 = \begin{pmatrix} a^2 & -4 & 4 \\ 0 & a^2 & 0 \\ 0 & 0 & a^2 \end{pmatrix}.$$

Die Minimalgleichung lautet $\psi(\lambda) = (a^2 - \lambda)^2$. Für $\alpha \neq 0$ ist die Gleichung lösbar (Satz 4). Wir transformieren mit

$$S = \begin{pmatrix} \frac{1}{4} & 2 & -1 \\ 0 & -1 & 1 \\ 0 & 3 & 1 \end{pmatrix}, \quad S^{-1} = \begin{pmatrix} 4 & 5 & -1 \\ 0 & -\frac{1}{4} & \frac{1}{4} \\ 0 & \frac{3}{4} & \frac{1}{4} \end{pmatrix}$$

[8]) Vgl. F. Cecioni [Rendiconti Acad. Lincei Roma (5) 18^1 (1909), 566—571]. G. Frobenius [Sitzungsberichte Akad. Berlin 1910, 3—15].

und erhalten

$$(SXS^{-1})^2 = Y^2 = \begin{pmatrix} a^2 & 1 & 0 \\ 0 & a^2 & 0 \\ 0 & 0 & a^2 \end{pmatrix}.$$

Diese Gleichung zerfällt in

$$Y_1^2 = \begin{pmatrix} a^2 & 1 \\ 0 & a^2 \end{pmatrix}; \quad Y_2^2 = a^2.$$

Als Lösungen erhält man:

$$Y_1^{(1)} = \begin{pmatrix} a & \frac{1}{2a} \\ 0 & a \end{pmatrix}; \quad Y_1^{(2)} = \begin{pmatrix} -a & -\frac{1}{2a} \\ 0 & -a \end{pmatrix}; \quad Y_2^{(1)} = a; \quad Y_2^{(2)} = -a$$

und daher

$$Y^{(1)} = \begin{pmatrix} a & \frac{1}{2a} & 0 \\ 0 & a & 0 \\ 0 & 0 & a \end{pmatrix}; \quad Y^{(2)} = \begin{pmatrix} a & \frac{1}{2a} & 0 \\ 0 & a & 0 \\ 0 & 0 & -a \end{pmatrix};$$

$$Y^{(3)} = \begin{pmatrix} -a & -\frac{1}{2a} & 0 \\ 0 & -a & 0 \\ 0 & 0 & -a \end{pmatrix}; \quad Y^{(4)} = \begin{pmatrix} -a & -\frac{1}{2a} & 0 \\ 0 & -a & 0 \\ 0 & 0 & a \end{pmatrix}.$$

Daraus $X^{(i)} = S^{-1} Y^{(i)} S$ d.h.:

$$X^{(1)} = \begin{pmatrix} a & -\frac{2}{a} & \frac{2}{a} \\ 0 & a & 0 \\ 0 & 0 & a \end{pmatrix}; \quad X^{(2)} = \begin{pmatrix} a & 6a - \frac{2}{a} & 2a + \frac{2}{a} \\ 0 & -\frac{a}{2} & -\frac{a}{2} \\ 0 & -\frac{3}{2}a & \frac{a}{2} \end{pmatrix};$$

$$X^{(3)} = \begin{pmatrix} -a & \frac{2}{a} & -\frac{2}{a} \\ 0 & -a & 0 \\ 0 & 0 & -a \end{pmatrix} = -X^{(1)};$$

$$X^{(4)} = \begin{pmatrix} -a & \frac{2}{a} - 6a & -\frac{2}{a} - 2a \\ 0 & \frac{a}{2} & \frac{a}{2} \\ 0 & \frac{3}{2}a & -\frac{a}{2} \end{pmatrix} = -X^{(2)}.$$

Hier lassen sich die Matrizen $Y^{(i)}$ nur auf eine Art in $\begin{pmatrix} Z_1 & O \\ O & Z_2 \end{pmatrix}$ zerlegen, nämlich so, daß Z_1 eine quadratische Matrix zweiter und Z_2 erster Ordnung ist. Aus $Y^{(2)}$ und $Y^{(4)}$ erhalten wir als $\begin{pmatrix} Z_1 & U \\ O & Z_2 \end{pmatrix}$ bzw. $\begin{pmatrix} Z_1 & O \\ V & Z_2 \end{pmatrix}$:

$$Z^{(2_1)}=\begin{pmatrix} a & \frac{1}{2a} & -\frac{1}{2a} \\ 0 & a & 0 \\ 0 & 0 & -a \end{pmatrix}; \quad Z^{(2_2)}=\begin{pmatrix} a & \frac{1}{2a} & 0 \\ 0 & a & 0 \\ 0 & -\frac{1}{2a} & -a \end{pmatrix}$$

und $Z^{(4_1)}=-Z^{(2_1)}$; $Z^{(4_2)}=-Z^{(2_2)}$,

und daher als weitere Lösungen der vorgelegten Gleichung:

$$X^{(2_1)}=S^{-1}Z^{(2_1)}S=\begin{pmatrix} a & 6a-\frac{8}{a} & 2a \\ 0 & -\frac{a}{2} & -\frac{a}{2} \\ 0 & -\frac{3a}{2} & \frac{a}{2} \end{pmatrix}; \quad X^{(4_1)}=-X^{(2_1)};$$

$$X^{(2_2)}=S^{-1}Z^{(2_2)}S=\begin{pmatrix} a & 6a-\frac{5}{2a} & 2a+\frac{5}{2a} \\ 0 & -\frac{a}{2}+\frac{1}{8a} & -\frac{a}{2}-\frac{1}{8a} \\ 0 & -\frac{3}{2}a+\frac{1}{8a} & \frac{a}{2}-\frac{1}{8a} \end{pmatrix}; \quad X^{(4_2)}=-X^{(2_2)}.$$

Aus $Y^{(1)}$ und $Y^{(2)}$ resultieren keine weiteren Lösungen, da $U=V=0$ ist.

B) Als Anwendung von Satz 5 behandeln wir das Beispiel (vgl. W. E. Roth [1])

$$X^2=A=\begin{pmatrix} 1 & -8 & 0 & 8 \\ 1 & 1 & 0 & -1 \\ 4 & 0 & 4 & -4 \\ 1 & -8 & 0 & 8 \end{pmatrix}.$$

A ist singulär, Minimalgleichung ist $\psi(\lambda)=\lambda^4-14\lambda^3+49\lambda^2-36\lambda=0$ mit den Wurzeln $\alpha_1=0$; $\alpha_2=1$; $\alpha_3=4$; $\alpha_4=9$. Wir zerlegen den Restklassenring; es ist

$$\pi_1(\lambda)=(\lambda-1)(\lambda-4)(\lambda-9)=\lambda^3-14\lambda^2+49\lambda-36;$$
$$\pi_2(\lambda)=\lambda(\lambda-4)(\lambda-9)=\lambda^3-13\lambda^2+36\lambda;$$
$$\pi_3(\lambda)=\lambda(\lambda-1)(\lambda-9)=\lambda^3-10\lambda^2+9\lambda;$$
$$\pi_4(\lambda)=\lambda(\lambda-1)(\lambda-4)=\lambda^3-5\lambda^2+4\lambda.$$

Die Polynome $\varrho_i(\lambda)$, so daß $\sum_{j=1}^{\nu}\varrho_j(\lambda)\pi_j(\lambda)=1$ ist, finden wir auf die folgende einfache Weise:

Es ist

$$\pi_1-\pi_2=-\lambda^2+13\lambda-36,$$
$$\pi_3-\pi_4=-5\lambda^2+5\lambda,$$

daher

$$5\pi_1-5\pi_2-\pi_3+\pi_4=60\lambda-5\cdot 36.$$

Ferner:

$$\pi_1 - \pi_3 = -4\lambda^2+40\lambda-36,$$
$$-4\pi_1+4\pi_2 = 4\lambda^2-4\cdot 13\lambda+4\cdot 36,$$

daher $$-3\pi_1+4\pi_2-\pi_3 = -12\lambda+3\cdot 36$$

oder $$-15\pi_1+20\pi_2-5\pi_3 = -60\lambda+15\cdot 36.$$

Folglich hat man: $-\frac{1}{36}\pi_1(\lambda)+\frac{1}{24}\pi_2(\lambda)-\frac{1}{60}\pi_3(\lambda)+\frac{1}{360}\pi_4(\lambda)=1.$

Die Polynome $\varrho_i(\lambda)$ sind Konstante. Der letzten Relation entsprechend hat man:

$$e_1(\lambda) = -\frac{1}{36}(\lambda^3-14\lambda^2+49\lambda-36);\ e_2(\lambda)=\frac{1}{24}(\lambda^3-13\lambda^2+36\lambda),$$
$$e_3(\lambda) = -\frac{1}{60}(\lambda^3-10\lambda^2+9\lambda);\ e_4(\lambda)=\frac{1}{360}(\lambda^3-5\lambda^2+4\lambda).$$

Lösungen unserer Gleichung erhalten wir, indem wir bilden:

$$x(\lambda) = \sum_{i=1}^{4} e_i(\lambda)\xi_i(\lambda).$$

Wir hatten $\xi_1=0;\ \xi_2=\pm 1;\ \xi_3=\pm 2;\ \xi_4=\pm 3$ und wählen etwa die Kombination: $\xi_1=0;\ \xi_2=1;\ \xi_3=2;\ \xi_4=-3$. Dann ist

$$x(\lambda) = -\frac{1}{6}\lambda^2+\frac{7}{6}\lambda,$$

infolgedessen haben wir:

$$X_1 = -\frac{1}{6}(A^2-7A)$$

als eine Lösung. Ihr negativer Wert ist ebenfalls eine Lösung. Wir erhalten sie, indem wir $\xi_1=0;\ \xi_2=-1;\ \xi_3=-2;\ \xi_4=3$ setzen. Entsprechend ergeben sich die anderen Lösungen.

C) Eine einfache, vom Vorhergehenden unabhängige Untersuchung der allgemeineren Polynome $P(X, A)$ in R auf Nullstellen ermöglicht der Umstand, daß zwei Elemente aus $K[\lambda;\ \mathrm{mod}\ \psi(\lambda)]$ dann und nur dann zueinander kongruent sind, wenn die entsprechenden Koeffizienten übereinstimmen. Denn aus

$$c_0+c_1\lambda+\ldots+c_{\nu-1}\lambda^{\nu-1} \equiv d_0+d_1\lambda+\ldots+d_{\nu-1}\lambda^{\nu-1} \bmod \psi(\lambda)$$

folgt

$$g(\lambda) = (c_0-d_0)+\ldots+(c_{\nu-1}-d_{\nu-1})\lambda^{\nu-1} \equiv 0 \bmod \psi(\lambda).$$

Diese Kongruenz ist aber dann und nur dann möglich, wenn $g(\lambda)$ identisch Null, also $c_i=d_i$ ist.

Jeder Kongruenz entspricht somit ein System von ν Gleichungen.

Auf die Resultate, die sich aus dieser einfachen Feststellung ergeben, soll nicht näher eingegangen werden. Wir wollen uns nur auf diese Äquivalenz zwischen einer Matrizengleichung und

einem System von Gleichungen stützen, um das folgende Beispiel in einfacher Weise zu behandeln:

$$X^2-3X=\begin{pmatrix}-3 & 3 & 6\\ -\frac{7}{3} & 5 & 2\\ -\frac{4}{3} & 4 & 0\end{pmatrix}.$$

Das Polynom niedrigsten Grades ist $\psi(\lambda)=\lambda^3-2\lambda^2-8\lambda$. Demnach hat man $\alpha_1=0$; $\alpha_2=-2$; $\alpha_3=4$. Nach Satz 4 ist diese Gleichung lösbar. Nach Zerlegung des Restklassenrings $K[\lambda;\ \mathrm{mod}\ \lambda(\lambda+2)(\lambda-4)]$ könnte man die Lösungen leicht finden. Wir wollen die Wurzeln der Matrizengleichung nach § 5c ermitteln.

Wir haben, wenn $\xi_0+\xi_1A+\xi_2A^2=\xi(A)$ eine Wurzel ist, die Kongruenz

$$(\xi(\lambda))^2-3\xi(\lambda)\equiv\lambda \bmod \lambda^3-2\lambda^2-8\lambda$$

zu lösen. Das der Kongruenz äquivalente Gleichungssystem ist:

$$\begin{aligned}\xi_1^2+2\xi_0\xi_1+4\xi_1\xi_2+12\xi_2^2&=3\xi_2\\ 2\xi_0\xi_1+16\xi_1\xi_2+16\xi_2^2&=3\xi_1+1\\ \xi_0^2&=3\xi_0\,.\end{aligned}$$

Die Lösungen sind:

$$\begin{aligned}(\xi_0;\ \xi_1;\ \xi_2)=&\left(0;\ -\tfrac{5}{12};\ \tfrac{1}{24}\right);\ \left(3;\ \tfrac{5}{12};\ -\tfrac{1}{24}\right);\\ &\left(0;\ 0;\ \tfrac{1}{4}\right);\ \left(3;\ 0;\ -\tfrac{1}{4}\right);\\ &\left(0;\ -\tfrac{1}{3};\ \tfrac{1}{3}\right);\ \left(3;\ \tfrac{1}{3};\ -\tfrac{1}{3}\right);\\ &\left(0;\ -\tfrac{3}{4};\ \tfrac{1}{8}\right);\ \left(3;\ \tfrac{3}{4};\ -\tfrac{1}{8}\right).\end{aligned}$$

Dementsprechend erhält man als Lösungen für die Matrizengleichung:

$$X_1=\tfrac{1}{24}\cdot(-10A+A^2);\quad X_2=\tfrac{1}{4}A^2;\quad X_3=\tfrac{1}{3}(-A+A^2);$$
$$X_4=\tfrac{1}{8}(-6A+A^2);\quad X_5=-\tfrac{1}{24}(-72E-10A+A^2);$$
$$X_6=-\tfrac{1}{4}(-12E+A^2);\quad X_7=-\tfrac{1}{3}(-9E-A+A^2);$$
$$X_8=-\tfrac{1}{8}(-24E-6A+A^2).$$

(Eingegangen den 1. November 1933.)