Miroslav Šisler

Über die Konvergenzbeschleunigung komplexer Iterationsverfahren

Aplikace matematiky, Vol. 15 (1970), No. 3, 156–176

Persistent URL: http://dml.cz/dmlcz/103283

ÜBER DIE KONVERGENZBESCHLEUNIGUNG KOMPLEXER ITERATIONSVERFAHREN

MIROSLAV ŠISLER

(Eingegangen am 21. November 1968)

In den Artikeln [2] und [3] wurde eine gewisse Methode für die Konvergenzbeschleunigung bei der iterativen Lösung eines Systems linearer algebraischer Gleichungen (siehe auch [1]) untersucht. Es handelt sich um die folgende Methode: Es sei das System m algebraischer Gleichungen $\mathbf{A}\mathbf{x} = \mathbf{b}$ gegeben und das Iterationsverfahren sei durch die Formel

$$\mathbf{x}_{\nu+1} = \mathbf{P}_k^{-1}\mathbf{Q}_k\mathbf{x}_\nu + \mathbf{P}_k^{-1}\mathbf{b}\,, \quad \nu = 0, 1, 2, \ldots \tag{1}$$

definiert, wo $\mathbf{P}_k = k\mathbf{P}_1$, $\mathbf{Q}_k = (k-1)\,\mathbf{P}_1 + \mathbf{Q}_1$, $k \neq 0$ ist. Dabei ist $\mathbf{A} = \mathbf{P}_1 - \mathbf{Q}_1$ eine solche Zerlegung der Matrix $\mathbf{A}$, dass der Spektralradius $\varrho(\mathbf{P}_1^{-1}\mathbf{Q}_1)$ der Matrix $\mathbf{P}_1^{-1}\mathbf{Q}_1$ kleiner also 1 ist, d.h. dass die durch die Formel (1) gegebene Methode für $k = 1$ konvergiert.

Die Artikel [2] und [3] befassen sich mit der Optimierung des Spektralradius $\varrho(\mathbf{P}_k^{-1}\mathbf{Q}_k)$ bezüglich des Parameters k, und zwar nicht nur im Falle des reellen Spektrums der Matrix $\mathbf{P}_1^{-1}\mathbf{Q}_1$, sondern auch dann, wenn die Matrix $\mathbf{P}_1^{-1}\mathbf{Q}_1$ ein komplexes Spektrum hat; dabei werden nur reelle Werte k zugelassen. In diesem Artikel sucht man aber den optimalen Parameter k im komplexen Gebiete. Das ist gerade bei komplexen Iterationsverfahren von grosser Bedeutung, da man durch die Wahl eines komplexen optimalen Parameters eine noch schnellere Konvergenz als für einen beliebigen reellen Wert k erzielen kann.

Man bezeichne jetzt mit $\mu_i(k)$ einen Eigenwert der Matrix $\mathbf{P}_k^{-1}\mathbf{Q}_k$ und mit $\lambda_i = \mu_i(1)$ einen Eigenwert der Matrix $\mathbf{P}_1^{-1}\mathbf{Q}_1$. Es seien λ_i, $i = 1, \ldots, n$ alle paarweise verschiedene Eigenwerte der Matrix $\mathbf{P}_1^{-1}\mathbf{Q}_1$. Es wurde im Artikel [2] (Satz 2.1) bewiesen, dass die Eigenwerte der Matrix $\mathbf{P}_k^{-1}\mathbf{Q}_k$ von k nicht abhängen und dass für die zu gleichem Eigenvektor entsprechenden Eigenwerte $\mu_i(k)$ und λ_i die Beziehung

$$\mu_i(k) = \frac{\lambda_i - 1}{k} + 1\,, \quad i = 1, \ldots, n \tag{2}$$

gilt.

Wir führen jetzt die Bezeichnung $\Lambda_i = \lambda_i - 1$ ein. Da man $\varrho(\mathbf{P}_1^{-1}\mathbf{Q}_1) < 1$ voraussetzt, ist $|\lambda_i| < 1$ für alle Zahlen i und die voneinander verschiedene Zahlen Λ_i liegen dann innerhalb des Kreises mit dem Radius 1 und dem Mittelpunkt im Punkte -1. Es ist also immer $\Lambda_i \neq 0$.

Für $k \neq 0$ definieren wir nun die Funktionen g_i der komplexen Veränderlichen $k = k_1 + k_2\mathrm{i}$ mit Hilfe der Formel

$$(3) \qquad g_i(k) = |\mu_i(k)|^2 = \frac{|\Lambda_i|^2}{k_1^2 + k_2^2} + \frac{2(R_ik_1 + J_ik_2)}{k_1^2 + k_2^2} + 1\,, \quad i = 1, \ldots, n\,,$$

wo $\Lambda_i = R_i + \mathrm{i}J_i$ ist.

Bemerkung 1. Manche Sätze in diesem Artikel gelten unabhängig von der Voraussetzung $\varrho(\mathbf{P}_1^{-1}\mathbf{Q}_1) < 1$. Es handelt sich um die Sätze 1, 2, 3, 5, 6, 7, 9, 11 und um die Bemerkung 3. Bei übrigen Sätzen führt man diese Voraussetzung ausdrücklich an.

1. *Die Funktionen g_i haben die folgenden Eigenschaften:*

a) *sie sind für jedes $k \neq 0$ definiert;*

b) *für jedes $k \neq 0$ gilt $g_i(k) \geqq 0$;*

c) *es gilt $g_i(k) < 1$ genau dann, wenn k in der Halbebene*

$$2R_ik_1 + 2J_ik_2 + |\Lambda_i|^2 < 0$$

liegt (offensichtlich liegt der Koordinatenursprung ausserhalb dieser Halbebene);

d) *die Funktionen g_i besitzen für $k \neq 0$ stetige erste partielle Ableitungen; es existiert ein einziger Punkt $k = -\Lambda_i$, in dem $\partial g_i/\partial k_1 = \partial g_i/\partial k_2 = 0$ ist, und die Funktionen g_i nehmen hier ihr absolutes Minimum an, das gleich Null ist.*

Beweis. Die Eigenschaften a), b) sind evident. Die Behauptung c) bekommt man sofort mit Hilfe der Formel (3). Die Behauptung d) beweist man wie folgt: Legt man die Ableitungen $\partial g_i/\partial k_1$, $\partial g_i/\partial k_2$ gleich Null, bekommt man die Gleichungen

$$-k_1|\Lambda_i|^2 + R_i(k_1^2 + k_2^2) - 2(R_ik_1 + J_ik_2)\,k_1 = 0\,,$$
$$-k_2|\Lambda_i|^2 + J_i(k_1^2 + k_2^2) - 2(R_ik_1 + J_ik_2)\,k_2 = 0\,.$$

Aus diesen Gleichungen bekommt man nach einer leichten Umformung die Beziehungen

$$|\Lambda_i|^2 + (R_ik_1 + J_ik_2) = 0\,,$$
$$[|\Lambda_i|^2 + 2(R_ik_1 + J_ik_2)]\,(J_ik_1 - R_ik_2) = 0\,.$$

Hiervon folgen sofort die Gleichungen $R_i k_1 + J_i k_2 = -|A_i|^2$ und $J_i k_1 - R_i k_2 = 0$, die nur eine einzige Lösung $k_1 = -R_i$, $k_2 = -J_i$, d.h. $k = -A_i$ besitzen. Es ist offensichtlich $g_i(-A_i) = 0$ und die Behauptung d) ist bewiesen.

Den Zusammenhang zwischen den Funktionen g_i, g_j, behandelt der folgende Hilfssatz:

2. *Ist $i \neq j$, dann ist die Menge der Punkte k, für die $g_i(k) = g_j(k)$ gilt, eine durch die Gleichung*

$$p_{ij} \equiv \alpha_{ij} k_1 + \beta_{ij} k_2 + \gamma_{ij} = 0$$

beschriebene Gerade p_{ij}. Dabei bezeichnet $\alpha_{ij} = R_i - R_j$, $\beta_{ij} = J_i - J_j$, $\gamma_{ij} = \frac{1}{2}(|A_i|^2 - |A_j|^2)$.

Der Beweis ist evident.

Nun bestimmen wir auf den Geraden p_{ij} diejenigen Punkte, in denen die Funktion g_i (bzw. g_j) ihren auf p_{ij} gebundenen Minimalwert annimmt. Es gilt dieser Hilfssatz:

3. I. *Es sei $|A_i| \neq |A_j|$. Dann hat die Funktion g_i (bzw. g_j) die folgenden Eigenschaften:*

a) *Die Funktion g_i (bzw. g_j) nimmt auf der Geraden p_{ij} ihren einzigen (lokalen und gleichzeitig absoluten) Minimalwert im Punkte $k^{ij} = k_1^{ij} + \mathrm{i}k_2^{ij}$ an, wobei*

$$k_1^{ij} = \frac{-(R_i + R_j)|A_i|\,|A_j| - R_j|A_i|^2 - R_i|A_j|^2}{2|A_i|\,|A_j| + 2(R_i R_j + J_i J_j)}, \tag{4}$$

$$k_2^{ij} = \frac{-(J_i + J_j)|A_i|\,|A_j| - J_j|A_i|^2 - J_i|A_j|^2}{2|A_i|\,|A_j| + 2(R_i R_j + J_i J_j)} \tag{5}$$

ist.

Dabei gilt

$$g_i(k^{ij}) = g_j(k^{ij}) = \frac{|A_i - A_j|^2}{(|A_i| + |A_j|)^2} < 1\,. \tag{6}$$

b) *Wenn keine reelle Zahl p existiert, für die $A_j = pA_i$ gilt, nimmt die Funktion g_i (bzw. g_j) auf der Geraden p_{ij} ihren (lokalen und gleichzeitig absoluten) Maximalwert im Punkte $l^{ij} = l_1^{ij} + \mathrm{i}l_2^{ij}$ an, wobei*

$$l_1^{ij} = \frac{-(R_i + R_j)|A_i|\,|A_j| + R_j|A_i|^2 + R_i|A_j|^2}{2|A_i|\,|A_j| - 2(R_i R_j + J_i J_j)}, \tag{4'}$$

$$l_2^{ij} = \frac{-(J_i + J_j)|A_i|\,|A_j| + J_j|A_i|^2 + J_i|A_j|^2}{2|A_i|\,|A_j| - 2(R_i R_j + J_i J_j)} \tag{5'}$$

ist. Dabei gilt

$$g_i(l^{ij}) = g_j(l^{ij}) = \frac{|A_i - A_j|^2}{(|A_i| - |A_j|)^2} > 1\,. \tag{6'}$$

Die Punkte k^{ij}, l^{ij} teilen dann die Gerade p_{ij} in drei Abschnitte, in denen die Funktion g_i (bzw. g_j) echt monoton ist und $\lim_{k\to\infty} g_i(k) = 1$ ist. Der Verlauf der Funktion g_i auf der Geraden p_{ij} ist in Abb. 1 angedeutet.

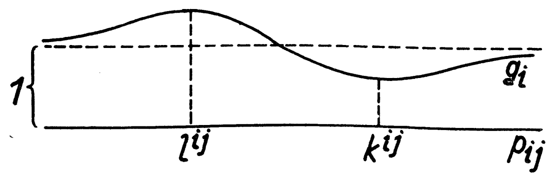

Abb. 1.

c) *Wenn eine reelle Zahl $p > 0$ mit $A_j = pA_i$ existiert, teilt der Punkt k^{ij} die Gerade p_{ij} in zwei Abschnitte, in denen die Funktion g_i (bzw. g_j) echt monoton ist, und es gilt wieder $\lim_{k\to\infty} g_i(k) = 1$. Der Verlauf der Funktion g_i auf der Geraden p_{ij} ist in diesem Falle in Abb. 2 angedeutet.*

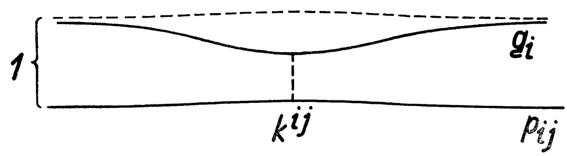

Abb. 2.

II. *Es sei $|A_i| = |A_j|$, $A_j \neq -A_i$. Dann enthält die Gerade p_{ij} den Punkt $k = 0$ und die Funktion g_i besitzt einen einzigen lokalen und gleichzeitig absoluten Minimalwert im Punkte $k^{ij} = k_1^{ij} + \mathrm{i}k_2^{ij}$, wo*

$$k_1^{ij} = \frac{-(R_i + R_j)\,|A_i|^2}{|A_i|^2 + (R_iR_j + J_iJ_j)}, \tag{4''}$$

$$k_2^{ij} = \frac{-(J_i + J_j)\,|A_i|^2}{|A_i|^2 + (R_iR_j + J_iJ_j)} \tag{5''}$$

ist. Dabei ist

$$g_i(k^{ij}) = g_j(k^{ij}) = \frac{|A_i - A_j|^2}{4|A_i|^2} < 1\,. \tag{6''}$$

Die Punkte 0, k^{ij} *teilen dann die Gerade* p_{ij} *in drei Abschnitte, in denen die Funktion* g_i (*bzw.* g_j) *echt monoton ist und es gilt* $\lim_{k\to\infty} g_i(k) = 1$, $\lim_{k\to 0} g_i(k) = +\infty$. *Der Verlauf der Funktion* g_i *auf der Geraden* p_{ij} *ist in Abb.* 3 *angedeutet.*

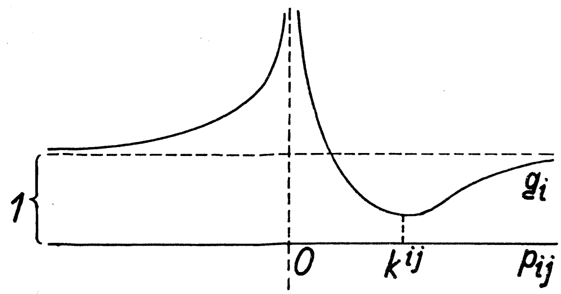

Abb. 3.

Beweis. Die Gleichung der Geraden p_{ij} kann man in der Form

$$Ak - \bar{A}\bar{k} + C = 0 \tag{7}$$

schreiben, wo $A = \beta_{ij} + i\alpha_{ij}$, $C = 2i\gamma_{ij}$ ist. Man untersuche nun die Menge der Punkte in der komplexen Ebene, für welche die Beziehung

$$u = \frac{A_i}{k} + 1 \tag{8}$$

gilt, wo k die Punkte der Geraden (7) sind. Wenn man $k = A_i/(u - 1)$ in die Gleichung der Geraden (7) einsetzt, bekommt man die Gleichung

$$C\bar{u}u - (\bar{B} + C)\,u + (B - C)\,\bar{u} + C - B + \bar{B} = 0\,, \tag{9}$$

wo $B = AA_i$ ist.

Nun unterscheidet man zwischen zwei Fällen.

I. $C \neq 0$, d.h. $|A_i| \neq |A_j|$. Dann entspricht der Geraden (7) in der durch die Formel (8) definierten Kreisinversion eine Kreislinie von der Gleichung (9). Es ist klar, dass diese Kreislinie immer durch den Punkt $u = 1$ geht, der dem Punkte $k = \infty$ auf der Geraden (7) entspricht. Da man jetzt solche Punkte k sucht, für die die Zahl $|A_i/k + 1|^2 = |u|^2$ minimal (bzw. maximal) ist, muss man also einen solchen Punkt $u = u^{ij}$ auf der Kreislinie (9) suchen, der den minimalen Abstand vom Punkt $u = 0$ hat, und ferner einen solchen Punkt $u = v^{ij}$, der den maximalen Abstand vom Punkt $u = 0$ hat. Die gesuchten Punkte u sind also die Schnittpunkte der durch den Punkt 0 gehenden Normale mit der Kreislinie (9) (siehe Abb. 4).

Setzt man $\bar{D} = B/C - 1$, $E = 1 + (\bar{B} - B)/C$ (E ist offensichtlich reell), nimmt die Gleichung (9) die Form

$$(10) \qquad u\bar{u} + Du + \bar{D}\bar{u} + E = 0$$

an. Setzt man $D = d_1 + d_2 i$, bekommt man aus (10) nach einer Umformung die Beziehung

$$(11) \qquad (u_1 + d_1)^2 + (u_2 - d_2)^2 = (d_1 + 1)^2 + d_2^2 .$$

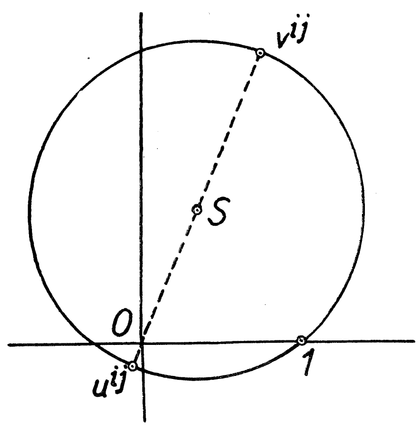

Abb. 4.

Nach einer leichten Berechnung stellt man fest, dass die Normale zu der Kreislinie (9), die durch den Punkt 0 geht, die Kreislinie (9) in den Punkten $u^{ij} = u_1^{ij} + iu_2^{ij}$ und $v^{ij} = v_1^{ij} + iv_2^{ij}$ schneidet, wobei

$$u_1^{ij} = -d_1\left[1 - \sqrt{\left(1 + \frac{2d_1 + 1}{d_1^2 + d_2^2}\right)}\right],$$

$$u_2^{ij} = \quad d_2\left[1 - \sqrt{\left(1 + \frac{2d_1 + 1}{d_1^2 + d_2^2}\right)}\right],$$

$$v_1^{ij} = -d_1\left[1 + \sqrt{\left(1 + \frac{2d_1 + 1}{d_1^2 + d_2^2}\right)}\right],$$

$$v_2^{ij} = \quad d_2\left[1 + \sqrt{\left(1 + \frac{2d_1 + 1}{d_1^2 + d_2^2}\right)}\right]$$

ist. Aus der Beziehung $\bar{D} = B/C - 1$ folgen jetzt die Beziehungen

$$d_1 = \frac{\alpha_{ij}R_i + \beta_{ij}J_i - 2\gamma_{ij}}{2\gamma_{ij}},$$

$$d_2 = \frac{\beta_{ij}R_i - \alpha_{ij}J_i}{2\gamma_{ij}},$$

wo α_{ij}, β_{ij}, γ_{ij} die in dem Hilfssatze 2 definierten Zahlen sind. Nach Einsetzung für d_1, d_2 in die vorhergehenden Formeln bekommt man nach einer Umformung die Beziehungen

$$u_1^{ij} = \frac{|A_j|^2 - (R_iR_j + J_iJ_j)}{(|A_i| + |A_j|)\,|A_j|}, \tag{12}$$

$$u_2^{ij} = \frac{R_iJ_j - R_jJ_i}{(|A_i| + |A_j|)\,|A_j|}, \tag{13}$$

$$v_1^{ij} = \frac{-|A_j|^2 + (R_iR_j + J_iJ_j)}{(|A_i| - |A_j|)\,|A_j|}, \tag{12'}$$

$$v_2^{ij} = \frac{-R_iJ_j + R_jJ_i}{(|A_i| - |A_j|)\,|A_j|}. \tag{13'}$$

Es ist nicht schwer zu beweisen, dass der Punkt u^{ij} dem Punkt $u = 0$ immer näher liegt als der Punkt v^{ij}, d.h. dass der Punkt k^{ij} (bzw. l^{ij}), der nach der Beziehung (8) dem Punkte u^{ij} (bzw. v^{ij}) entspricht, der Minimalpunkt (bzw. Maximalpunkt) der Funktion g_i (bzw. g_j) auf der Geraden p_{ij} ist.

Wir werden jetzt zwischen zwei Fällen unterscheiden. Zuerst setzt man voraus, dass der Mittelpunkt der Kreislinie (9) nicht auf der reellen Achse liegt. Da $S \equiv$ $\equiv [-d_1, d_2]$ ist, setzt man also voraus, dass $d_2 \neq 0$, d.h.

$$\beta_{ij}R_i - \alpha_{ij}J_i = (J_i - J_j)\,R_i - (R_i - R_j)\,J_i = J_iR_j - R_iJ_j \neq 0$$

ist. Es ist also $\begin{vmatrix} R_i & J_i \\ R_j & J_j \end{vmatrix} \neq 0$, sodass keine reelle Zahl p existiert, für die $A_j = pA_i$ gilt. Da die Kreislinie (9) immer durch den Punkt $u = 1$ geht, kann in diesem Falle nicht $v^{ij} = 1$ gelten (und es ist dann auch $u^{ij} \neq 1$). Mit Hilfe der Beziehungen (12), (13), (12'), (13') und der Formeln $u^{ij} = A_i/k^{ij} + 1$, $v^{ij} = A_i/l^{ij} + 1$ bekommt man dann die Beziehungen (4), (5), (4'), (5'). Durch Einsetzung von k^{ij} bzw. l^{ij} in $g_i(k)$, bekommt man ferner die Beziehungen (6), (6').

Nun beweisen wir die Ungleichung

$$\frac{|A_i - A_j|^2}{(|A_i| + |A_j|)^2} < 1$$

(siehe (6)). Es sei

$$\frac{|A_i - A_j|^2}{(|A_i| + |A_j|)^2} \geqq 1 .$$

Dann folgen sukzessiv die Ungleichungen

$$|A_i - A_j|^2 \geqq (|A_i| + |A_j|)^2 ,$$
$$|A_i|^2 + |A_j|^2 - 2(R_iR_j + J_iJ_j) \geqq |A_i|^2 + |A_j|^2 + 2|A_i|\,|A_j| ,$$
$$-(R_iR_j + J_iJ_j) \geqq |A_i|\,|A_j| .$$

Ist $R_iR_j + J_iJ_j > 0$, bekommt man einen Widerspruch. Ist $R_iR_j + J_iJ_j \leqq 0$, dann bekommt man durch Quadrieren der vorhergehenden Ungleichung sukzessiv die Ungleichungen

$$R_i^2R_j^2 + 2R_iR_jJ_iJ_j + J_i^2J_j^2 \geqq (R_i^2 + J_i^2)(R_j^2 + J_j^2) ,$$
$$2R_iR_jJ_iJ_j \geqq R_i^2J_j^2 + R_j^2J_i^2 ,$$
$$0 \geqq (R_iJ_j + R_jJ_i)^2 .$$

Da nach der Voraussetzung $R_iJ_j + R_jJ_i \neq 0$ ist, bekommt man einen Widerspruch. Es gilt also die Ungleichung (6). Ähnlich beweist man auch die Ungleichung (6′).

Im Falle, wenn keine reelle Zahl p existiert, für die $A_j = pA_i$ gilt, tritt die folgende Situation ein: dem Abschnitt ∞, l^{ij} entspricht in der durch die Formel (8) definierten Kreisinversion das Segment 1, v^{ij}, dem Abschnitt l^{ij}, k^{ij} das Segment v^{ij}, u^{ij} und endlich dem Abschnitt k^{ij}, ∞ das Segment u^{ij}, 1. Da für die Punkte u der Kreislinie (9) die Beziehung $|u|^2 = g_i(k)$ gilt, wo $k = A_i/(u - 1)$ ist, verläuft die Funktion g_i auf der Geraden p_{ij} wie in Abb. 1.

Nun untersuchen wir den Fall, wenn der Mittelpunkt der Kreislinie (9) auf der reellen Achse liegt. Dies entspricht dem Falle, wenn für eine gewisse reelle Zahl p die Beziehung $A_j = pA_i$ gilt. Geometrisch entspricht diese Beziehung der Situation, wenn die Punkte A_i, A_j auf der durch den Punkt $k = 0$ gehende Geraden liegen. Für $p > 0$ es gilt dann $u^{ij} \neq 1$, $v^{ij} = 1$. Mit Hilfe der Formel $u^{ij} = A_i/k^{ij} + 1$ und der Formeln (12), (13) bekommt man wieder für den Minimalpunkt die Formeln (4), (5) und nach Einsetzung für k_{ij} in $g_i(k)$ wieder die Formel (6).

Jetzt beweisen wir, dass wieder die Ungleichung

$$\frac{|A_i - A_j|^2}{(|A_i| + |A_j|)^2} < 1$$

(siehe (6)) gilt. Man setze voraus, dass die Ungleichung

$$\frac{|A_i - A_j|^2}{(|A_i| + |A_j|)^2} \geqq 1$$

gilt. Nach der Formel $A_j = pA_i$, $p > 0$ und nach der vorhergehenden Ungleichung folgen sukzessiv die Ungleichungen

$$\frac{|A_i - pA_i|^2}{(|A_i| + p|A_i|)^2} \geqq 1 ,$$

$$\frac{(1 - p)^2}{(1 + p)^2} \geqq 1 ,$$

$$-p \geqq p .$$

Das ist aber ein Widerspruch, denn es ist $p > 0$. Es gilt also wieder die Ungleichung aus der Formel (6).

Da dem Punkte $v^{ij} = 1$ der Punkt $l^{ij} = \infty$ entspricht, erreicht die Funktion g_i ihren Maximalwert auf der Geraden p_{ij} nicht im endlichen Bereich. In der Kreisinversion entsprechen dann offensichtlich den beiden Segmenten u^{ij}, 1 die Halbgeraden ∞, k^{ij} und k^{ij}, ∞, auf welchen die Funktion g_i monoton ist, und der Verlauf der Funktion g_i auf der Geraden p_{ij} entspricht der Abb. 2.

II. Falls $|A_i| = |A_j|$, $A_i \neq -A_j$, d.h. $C = 0$ ist, nimmt die Gerade (7) die Form

(14) $$A k - \bar{A}\bar{k} = 0$$

an (sie geht also durch den Punkt $k = 0$). Nach Einsetzung für k in (14) (siehe (8)), bekommt man wieder eine Gerade

(15) $$-\bar{B}u + B\bar{u} - B + \bar{B} = 0 .$$

Diese Gerade geht offensichtlich durch den Punkt $u = 1$.

Ähnlich wie im Falle I berechnet man, dass die Koordinaten des dem Punkte $u = 0$ nächstliegenden Punktes $u = u^{ij}$ durch die Formeln

(12′) $$u_1^{ij} = \frac{|A_i|^2 - (R_i R_j + J_i J_j)}{2|A_i|^2},$$

(13′) $$u_2^{ij} = \frac{R_i J_j - J_i R_j}{2|A_i|^2}$$

gegeben sind. Es ist offensichtlich $u^{ij} \neq 1$. Wenn nämlich $u = 1$ wäre, würde $u_2^{ij} = 0$ gelten, d.h. $R_i J_j - J_i R_j = 0$, was nicht möglich ist, denn es ist $|A_i| = |A_j|$, $A_i \neq -A_j$. Nach den Formeln (12′), (13′) und der Formel $u^{ij} = A_i/k^{ij} + 1$ bekommt man die Beziehungen (4″), (5″), welche nur einen Spezialfall der Beziehungen (4), (5) darstellen.

Nun beweisen wir die Ungleichung (6″):

$$\frac{|A_i - A_j|^2}{4|A_i|^2} < 1 .$$

Es sei $|\Lambda_i - \Lambda_j|^2/4|\Lambda_i|^2 \geqq 1$. Dann folgt sukzessiv

$$|\Lambda_i - \Lambda_j|^2 \geqq 4|\Lambda_i|^2 ,$$
$$2|\Lambda_i|^2 - 2(R_iR_j + J_iJ_j) \geqq 4|\Lambda_i|^2 ,$$
$$-(R_iR_j + J_iJ_j) \geqq |\Lambda_i|^2 .$$

Falls $R_iR_j + J_iJ_j > 0$ ist, bekommt man einen Widerspruch. Falls $R_iR_j + J_iJ_j \leqq 0$ ist, bekommt man durch Quadrierung der vorhergehenden Ungleichung und mit Hilfe der Beziehung $|\Lambda_i| = |\Lambda_j|$ sukzessiv die Ungleichungen

$$R_i^2R_j^2 + 2R_iR_jJ_iJ_j + J_i^2J_j^2 \geqq (R_i^2 + J_i^2)(R_j^2 + J_j^2) ,$$
$$2R_iR_jJ_iJ_j \geqq R_i^2J_j^2 + J_i^2R_j^2 ,$$
$$0 \geqq (R_iJ_j + J_iR_j)^2 .$$

Das ist aber ein Widerspruch, denn die Gleichung $R_iJ_j + J_iR_j = 0$ kann mit Rücksicht auf die Beziehungen $|\Lambda_i| = |\Lambda_j|$, $\Lambda_i \neq -\Lambda_j$ nicht gelten.

Die Punkte u^{ij}, 1 auf der Geraden (14) teilen die Gerade (15) in drei Abschnitte. Dem Abschnitt ∞, u^{ij} entspricht in der Kreisinversion auf der Geraden (14) der Abschnitt 0, k^{ij}, dem Abschnitt u^{ij}, 1 entspricht der Abschnitt k^{ij}, ∞ und dem Abschnitt 1, ∞ entspricht der Abschnitt ∞, 0. Dabei ist klar, dass in diesen Abschnitten der Geraden (14) die Funktion g_i echt monoton ist und dass $\lim_{k\to\infty} g_i(k) = 1$ ist. Der Verlauf der Funktion g_i entspricht also der Abb. 3.

Dadurch ist der Hilfssatz 3 bewiesen.

Bemerkung 2. Wenn $\varrho(\mathbf{P}_1^{-1}\mathbf{Q}_1) < 1$ ist, liegen die Punkte Λ_i im Kreise mit dem Radius 1 und dem Mittelpunkt im Punkte -1, sodass der Satz 3 alle möglichen Fälle behandelt.

Von den Eigenschaften der Punkte k^{ij} spricht der folgende Satz:

4. *Falls $\varrho(\mathbf{P}_1^{-1}\mathbf{Q}_1) < 1$ ist, liegt der Minimalpunkt der Funktion g_i (bzw. g_j) (d.h. der Punkt k^{ij}) auf der Geraden p_{ij} stets innerhalb des Kreises mit dem Mittelpunkt* 1 *und mit dem Radius* 1.

Beweis. Die Punkte innerhalb des Kreises K genügen der Ungleichung

$$|k - 1|^2 < 1 . \tag{16}$$

Wir beweisen, dass der Punkt k^{ij} die Ungleichung (16) erfüllt. Dazu genügt zu beweisen, dass der Punkt $u^{ij} = \Lambda_i/k^{ij} + 1$ (siehe (8)) die Ungleichung

$$u^{ij}\bar{\Lambda}_i + \bar{u}^{ij}\Lambda_i - |\Lambda_i|^2 - 2R_i > 0 \tag{17}$$

erfüllt. Diese Ungleichung entsteht aus der Ungleichung (16) nach Einsetzung aus der Formel $u^{ij} = \Lambda_i/k^{ij} + 1$. Falls man in (17) aus den Formeln (12) und (13) einsetzt, bekommt man nach einigen Umformungen die Ungleichung

$$|\Lambda_i|(2R_j + |\Lambda_j|^2) + |\Lambda_j|(2R_i + |\Lambda_i|^2) < 0 . \tag{18}$$

Da die Zahlen λ_i, λ_j im Einheitskreis liegen, gilt jetzt für i sukzessiv

$$\begin{gathered} |\lambda_i|^2 < 1 , \\ |\Lambda_i + 1|^2 < 1 , \\ 2R_i + |\Lambda_i|^2 < 0 . \end{gathered}$$

Ähnlicherweise bekommt man auch die Ungleichung

$$2R_j + |\Lambda_j|^2 < 0 . \tag{20}$$

Aus (19) und (20) folgt sofort, dass die Ungleichung (18) gilt, was zu beweisen war.

Wir untersuchen jetzt die drei Funktionen g_i, g_j, g_k, und die Menge der Punkte k, für die sich die durch die Gleichungen $z = g_i(k)$, $z = g_j(k)$, $z = g_k(k)$ gegebenen Flächen (d.h. die Graphen der Funktionen g_i, g_j, g_k) schneiden. Wir wissen schon, dass die Mengen der Punkte k, für die sich die einzelnen Fläche wechselseitig schneiden, die Geraden p_{ij}, p_{jk}, p_{ki} sind. Wir untersuchen jetzt die Menge der gemeinsamen Punkte der Geraden p_{ij}, p_{jk}, p_{ki}. Diese Punkte bekommt man als die Lösung des Gleichungssystems

$$\begin{aligned} \alpha_{ij}k_1 + \beta_{ij}k_2 + \gamma_{ij} &= 0 , \\ \alpha_{jk}k_1 + \beta_{jk}k_2 + \gamma_{jk} &= 0 , \\ \alpha_{ki}k_1 + \beta_{ki}k_2 + \gamma_{ki} &= 0 . \end{aligned} \tag{21}$$

Diese Gleichungen sind linear abhängig, da die Determinante der erweiterten Matrix des Systems (21) gleich Null ist, wie man leicht berechnen kann. Es handelt sich also um ein Gleichungssystem mit den Unbekannten k_1, k_2 vom Rang nicht grösser als 2. Es gilt die folgende Behauptung:

5. *Es seien Λ_i, Λ_j, Λ_k die Punkte, welche auf keiner Geraden liegen. Dann haben die Geraden p_{ij}, p_{jk}, p_{ki} einen gemeinsamen Schnittpunkt k^{ijk}. Falls die Punkte Λ_i, Λ_j, Λ_k auf einer Geraden liegen, sind die Geraden p_{ij}, p_{jk}, p_{ki} wechselseitig parallel (oder identisch).*

Beweis. Wie wir schon wissen, ist der Rang des Systems (21) höchstens gleich 2. Man setze daher voraus, dass z.B. die dritte Gleichung eine Linearkombination der ersten zwei Gleichungen ist. Dann ist der Rang des Systems (21) genau dann gleich 2, wenn die Matrix

$$\left\| \begin{matrix} \alpha_{ij}, \beta_{ij} \\ \alpha_{jk}, \beta_{jk} \end{matrix} \right\| \tag{22}$$

regulär ist. Das tritt aber dann und nur dann ein, wenn die Punkte Λ_i, Λ_j, Λ_k zu keiner Gerade gehören. Würden nämlich diese Punkte auf einer Geraden liegen, würde eine reelle Zahl $\varrho \neq 0$ mit

$$\Lambda_i - \Lambda_j = \varrho(\Lambda_j - \Lambda_k) \tag{23}$$

existieren, d.h. es würde $\alpha_{ij} = \varrho\alpha_{jk}$, $\beta_{ij} = \varrho\beta_{jk}$ gelten und die Matrix (22) wäre singulär. Das ist aber ein Widerspruch. Das System (21) besitzt dann genau eine Lösung k^{ijk}, womit der este Teil der Behauptung bewiesen ist.

Setzt man jetzt voraus, dass die Punkte Λ_i, Λ_j, Λ_k auf einer Geraden liegen, dann ist, wie wir schon wissen, die Matrix (22) singulär und die Vektoren $(\alpha_{ij}, \beta_{ij})$, $(\alpha_{jk}, \beta_{jk})$ gehören zu der gleichen Richtung. Eine ähnliche Situation entsteht auch bei den Vektoren $(\alpha_{ij}, \beta_{ij})$, $(\alpha_{ki}, \beta_{ki})$. Alle Geraden p_{ij}, p_{jk}, p_{ki} sind also parallel (oder identisch).

Nun können wir schon versuchen, die Frage über die Lage des optimalen Parameters k in der komplexen Ebene zu beantworten. Für eine beliebige Zahl $k \neq 0$ ist

$$\varrho(\mathbf{P}_k^{-1}\mathbf{Q}_k) = \max_i \sqrt{[g_i(k)]} \,.$$

Als *Optimalparameter* werden wir jetzt eine solche Zahl k_0 bezeichnen, für die der Ausdruck $\max_i \sqrt{[g_i(k)]}$ sein Infimum bezüglich k erreicht (aus den Verlauf der Funktionen g_i ist klar, dass das Infimun in diesem Falle auch ein Minimum ist). Wir definieren jetzt den *optimalen Spektralradius* als den Spektralradius $\varrho(\mathbf{P}_{k_0}^{-1}\mathbf{Q}_{k_0})$. Man kann also schreiben

$$\varrho(\mathbf{P}_{k_0}^{-1}\mathbf{Q}_{k_0}) = \min_k \varrho(\mathbf{P}_k^{-1}\mathbf{Q}_k) = \min_k \max_i \sqrt{[g_i(k)]} \,. \tag{24}$$

Es gilt der folgende Satz:

6. a) *Man setze voraus, dass wenigstens zwei Eigenwerte der Matrix* $\mathbf{P}_1^{-1}\mathbf{Q}_1$ *voneinander verschieden sind, und bezeichne mit* M_2 *die Menge aller solchen Punkte* $k = k^{ij}$, *dass für alle* $l = 1, \ldots, n$ *die Ungleichung*

$$g_i(k^{ij}) = g_j(k^{ij}) \geqq g_l(k^{ij})$$

gilt. Ferner bezeichne man mit M_3 *die Menge aller solchen Punkte* $k = k^{ijk}$, *dass* Λ_i, Λ_j, Λ_k *zu keiner Geraden gehören und dass für alle* $l = 1, \ldots, n$ *die Ungleichung*

$$g_i(k^{ijk}) = g_j(k^{ijk}) = g_k(k^{ijk}) \geqq g_l(k^{ijk})$$

gilt. Dann gilt für den Optimalparameter die Beziehung

$$\varrho(\mathbf{P}_{k_0}^{-1}\mathbf{Q}_{k_0}) = \min_{k \in \mathrm{M}_2 \cup \mathrm{M}_3} \max_i \sqrt{[g_i(k)]} = \min_{\substack{k^{ij} \in \mathrm{M}_2 \\ k^{ijk} \in \mathrm{M}_3}} \{\sqrt{[g_i(k^{ij})]}, \sqrt{[g_i(k^{ijk})]}\} \,.$$

b) *Falls die Matrix* $\mathbf{P}_1^{-1}\mathbf{Q}_1$ *einen m-fachen Eigenwert* λ *besitzt, nimmt der Spektralradius* $\varrho(\mathbf{P}_k^{-1}\mathbf{Q}_k)$ *seinen Minimalwert, der gleich Null ist, für* $k_0 = -\Lambda$ *an, wo* $\Lambda = \lambda - 1$ *ist.*

Bemerkung 3. Die reellen und imaginären Teile der Zahlen k^{ij} sind durch die Formeln (4) und (5) gegeben und der Wert $g_i(k^{ij})$ ist durch die Formel (6) gegeben. Für den Real- und Imaginärteil der Zahlen k^{ijk} folgen aus dem System (21) die folgenden Formeln:

(21′)

$$k_1^{ijk} = \frac{\gamma_{jk}\beta_{ij} - \gamma_{ij}\beta_{jk}}{\alpha_{ij}\beta_{jk} - \alpha_{jk}\beta_{ij}} = -\frac{1}{2}\frac{|\Lambda_i|^2(J_j - J_k) + |\Lambda_j|^2(J_k - J_i) + |\Lambda_k|^2(J_i - J_j)}{R_i(J_j - J_k) + R_j(J_k - J_i) + R_k(J_i - J_j)},$$

$$k_2^{ijk} = \frac{-\alpha_{ij}\gamma_{jk} + \alpha_{jk}\gamma_{ij}}{\alpha_{ij}\beta_{jk} - \alpha_{jk}\beta_{ij}} = -\frac{1}{2}\frac{|\Lambda_i|^2(R_j - R_k) + |\Lambda_j|^2(R_k - R_i) + |\Lambda_k|^2(R_i - R_j)}{J_i(R_j - R_k) + J_j(R_k - R_i) + J_k(R_i - R_j)}.$$

Die Formel für $g_i(k^{ijk})$ werden wir wegen ihrer Kompliziertheit nicht anführen.

Beweis. Man setze voraus, dass die Matrix $\mathbf{P}_1^{-1}\mathbf{Q}_1$ nur einen einzigen m-fachen Eigenwert λ besitzt. Dann ist $\Lambda_1 = \Lambda_2 = \ldots = \Lambda_m = \Lambda$, $\Lambda = \lambda - 1$, und auch $g_1 = g_2 = \ldots = g_m = g$. Nach Hilfsatz 1, Behauptung d), gilt, dass

$$\min_k \max_{i=1,\ldots,m} \sqrt{[g_i(k)]} = \min_k \sqrt{[g(k)]} = \sqrt{[g(-\Lambda)]} = 0$$

ist. Dadurch ist die Behauptung b) des Satzes 6 bewiesen.

Man setze voraus, dass mindestens zwei von den Eigenwerten der Matrix $\mathbf{P}_1^{-1}\mathbf{Q}_1$ voneinander verschieden sind, d.h. $n \geqq 2$ ist. Dann ist für jedes $k \neq 0$

$$\max_{i=1,\ldots,n} g_i(k) > 0\,. \tag{25}$$

Wenn nämlich für irgendein k_0

$$\max_{i=1,\ldots,n} g_i(k_0) = 0$$

wäre, würde $g_i(k_0) = 0$ sein für $i = 1, \ldots, n$. Da nach Hilfssatz 1, Behauptung d), die Funktion g_i ihren einzigen absoluten Minimalwert, der gleich Null ist, im Punkte $-\Lambda_i$ annimmt, wäre dann

$$k_0 = -\Lambda_i\,, \quad i = 1, \ldots, n\,.$$

Das ist aber ein Widerspruch, denn mindestens zwei von der Zahlen Λ_i sind verschieden.

Man setze voraus, dass $\varrho(\mathbf{P}_k^{-1}\mathbf{Q}_k)$ seinen Minimalwert für $k = k_0$ erreicht, sodass die Beziehungen

$$\min_k \max_i \sqrt{[g_i(k)]} = \max_i \sqrt{[g_i(k_0)]} = \sqrt{[g_{i_0}(k_0)]} \tag{26}$$

gelten, wo i_0 irgendeine von den Zahlen $1, 2, \ldots, n$ ist. Nach (25) ist

$$\sqrt{[g_{i_0}(k_0)]} > 0\,. \tag{27}$$

Ferner setze man voraus, dass der Punkt k_0 zu keiner von der Geraden p_{ij} gehört. Dann kann man einen solchen Punkt $k_1 \neq k_0$ wählen (genügend nahe dem Punkte k_0), dass die Beziehung

$$\max_i \sqrt{[g_i(k_1)]} = \sqrt{[g_{i_0}(k_1)]} \tag{28}$$

und die Ungleichung

$$g_{i_0}(k_1) < g_{i_0}(k_0) \tag{29}$$

gilt (nach (27) und Hilfssatz 1, Behauptung c) ist nämlich $\partial g_{i_0}/\partial k_1(k_0) \neq 0$, $\partial g_{i_0}/\partial k_2(k_0) \neq 0$). Aus (28), (29) und (26) folgen jetzt die Ungleichungen

$$\min_k \max_i \sqrt{[g_i(k)]} \leqq \max_i \sqrt{[g_i(k_1)]} = \sqrt{[g_{i_0}(k_1)]} <$$
$$< \sqrt{[g_{i_0}(k_0)]} = \min_k \max_i \sqrt{[g_i(k)]}\,,$$

was aber ein Widerspruch ist. Dadurch ist bewiesen, dass der Optimalparameter zu irgendeiner von der Geraden p_{ij} gehört. Ganz analogisch beweist man, dass von den Punkten auf den Geraden p_{ij} nur die Punkte k^{ij} in Betracht kommen, für die die Funktion g_i (bzw. g_j) ihr Minimum annimmt, d.h. die durch die Formeln (4), (5) gegebenen Punkte, oder die Punkte k^{ijk}, in denen sich die Geraden p_{ij}, p_{jk}, p_{ki} schneiden.

7. *Der Parameter k_0, für welchen der Spektralradius $\varrho(\mathbf{P}_k^{-1}\mathbf{Q}_k)$ sein Minimum annimmt, liegt im kleinsten konvexen Vieleck, das die Punkte k^{ij}, $i, j = 1, \ldots, n$ enthält.*

Beweis. Aus dem Satz 6 folgt, dass man als Optimalparameter k_0 immer eine von den Zahlen k^{ij} oder k^{ijk} wählen kann. Im ersten Falle ist die Behaptung des Satzes evident. Man setze jetzt voraus, dass der Optimalparameter in irgendeinem Punkt $k_0 \neq k^{ij}$ $(i, j = 1, \ldots, n)$ liegt, in dem sich drei oder mehr Flächen schneiden. Man setze voraus, dass es sich um Flächen g_i, $i \in \mathrm{N}_1$ handelt, wobei N_1 eine Untermenge der Menge $\mathrm{N} = \{1, \ldots, n\}$ ist, die drei oder mehr Elemente enthält. Für jedes $i \in \mathrm{N}_1$ und $j \in \mathrm{N}_2 = \mathrm{N} - \mathrm{N}_1$ muss jetzt die folgende Umgebung gelten:

$$g_i(k_0) > g_j(k_0)\,.$$

Es sei Ω eine solche Umgebung des Punktes k_0, dass für jedes $k \in \Omega$ die Ungleichungen

$$g_i(k) > g_j(k)\,, \quad g_i(k) < 1$$

gelten, wobei $i \in \mathrm{N}_1$, $j \in \mathrm{N}_2$ beliebige Indices sind. Man untersuche jetzt die Geraden p_{rs}, wo r, $s \in \mathrm{N}_1$ ist. Diese Geraden haben einen gemeinsamen Schnittpunkt im Punkte k_0. Der Punkt k_0 teilt jede von diesen Geraden in zwei Halbgeraden p'_{rs}, p''_{rs}. Wenn wir im folgenden von der Geraden p_{rs} oder von der Halbgeraden p'_{rs}, p''_{rs} sprechen werden, werden wir immer nur jenen Teil dieser Geraden meinen, der in der Umgebung Ω liegt. Die Halbegrade p'_{rs} nennt man eine Halbgerade vom *ersten Typ*, wenn für jeden Punkt $k \neq k_0$ dieser Halbgeraden die Ungleichungen $g_i(k) \leqq g_r(k) = g_s(k)$, $i \in \mathrm{N}_1$ gelten (geometrisch bedeutet dies, dass die Schnittlinie der Flächen g_r, g_s aus der Obenansicht sichtbar sind). Die Halbgeraden, die nicht vom ersten Typ sind, nennt man Halbgeraden vom *zweiten Typ*.

Ohne Verlust an Allgemeinheit kann man jetzt voraussetzen, dass keine von der Geraden p_{ij}, i, $j \in \mathrm{N}_1$ aus zwei Halbgeraden vom ersten Typ zusammengesetzt ist. Wenn nämlich beide Halbgeraden p'_{ij}, p''_{ij} vom ersten Typ wären, dann müsste $k_0 = k^{ij}$ gelten (das folgt nämlich aus dem Hilfssatze 3, denn der Abschnitt der Geraden p_{ij}, wo $g_i(k) = g_j(k) < 1$ ist, besteht aus zwei durch den Punkt k^{ij} getrennten Abschnitten, in denen die Funktion g_i (bzw. g_j) echt monoton ist.

Nun werden wir beweisen, dass die Halbgeraden vom ersten Typ nie in einer abgeschlossenen Halbebene p liegen können, deren Grenzgerade durch den Punkt k_0 geht. Man wähle irgendeine durch den Punkt k_0 gehende Gerade, die von allen Geraden p_{ij}, i, $j \in \mathrm{N}_1$ verschieden ist, und setze voraus, dass alle Halbgeraden vom ersten Typ in einer durch die Gerade p begrenzten abgeschlossenen Halbebene liegen.

Nach dem, was oben behauptet wurde, muss die entgegengesetzte Halbebene nur Halbgeraden vom zweiten Typ enthalten. Es existiert also offensichtlich ein solcher Index $i_0 \in \mathrm{N}$, dass für jeden Punkt dieser Halbebene die Ungleichung $g_{i_0}(k) > g_j(k)$ für alle $j \in \mathrm{N}_1 - \{i_0\}$ gilt. Wählt man also irgendein $j_0 \in \mathrm{N}_1 - \{i_0\}$, dann durchläuft die Gerade $p_{i_0 j_0}$ beide Halbebenen und in beiden Halbebenen existieren also Punkte, für die $g_{i_0}(k) = g_{j_0}(k)$ gilt, was ein Widerspruch ist.

Wir beweisen ferner, dass der Punkt k^{ij} nicht auf einer Halbgeraden p'_{ij} vom ersten Typ liegen kann. Wäre es so, würden für jedes k' aus dem Segment k^{ij}, k_0 die Ungleichungen

$$g_i(k') = g_j(k') < g_i(k_0) = g_j(k_0)\,,$$
$$g_i(k') = g_j(k') \geqq g_l(k')\,, \quad l \in \mathrm{N}_1$$

gelten. Die erste Ungleichung folgt wieder leicht aus dem Hilfssatz 3, nach dem der Punkt k^{ij} den Abschnitt der Geraden p_{ij}, wo $g_i(k) = g_j(k) < 1$ gilt, in zwei Abschnitte teilt, in denen die Funktion g_i (bzw. g_j) echt monoton ist. Die zweite Ungleichung folgt aus der Tatsache, dass die Halbgerade p'_{ij} vom ersten Typ ist. Aus diesen zwei Ungleichungen folgt sofort ein Widerspruch mit der Voraussetzung, dass k_0 ein Optimalparameter ist.

Da also die Halbgeraden vom ersten Typ in keiner Halbebene liegen, deren Grenzlinie durch den Punkt k_0 geht, und da die entsprechenden Punkte k^{ij} zu den entgegengesetzten Halbgeraden gehören, liegt der Punkt k_0 im kleinsten konvexen Vieleck, das die Punkte k^{ij} enthält. Dadurch ist der Satz 7 bewiesen.

Aus Satz 7 und Hilfssatz 4 folgt sofort der folgende Satz:

8. *Wenn $\varrho(\mathbf{P}_1^{-1}\mathbf{Q}_1) < 1$, liegt der Parameter k_0, für welchen der Spektralradius $\varrho(\mathbf{P}_k^{-1}\mathbf{Q}_k)$ sein Minimum annimmt, immer innerhalb des Kreises K mit dem Mittelpunkt 1 und mit dem Radius 1, d.h. im Kreise $|k - 1| < 1$.*

Man kann sofort einen weiteren Satz formulieren, der das Gebiet beschränkt, in dem die zulässigen Parameter k liegen, d.h. die Parameter, für die das Iterationsverfahren (1) konvergiert.

9. *Die Menge aller Parameter k, für die $\varrho(\mathbf{P}_k^{-1}\mathbf{Q}_k) < 1$ ist, ist durch den Durchschnitt der Halbebenen*

$$2R_i k_1 + 2J_i k_2 + |\Lambda_i|^2 < 0\,, \quad i = 1, \ldots, n \tag{30}$$

gebildet.

Beweis. Es ist klar, dass aus der Beziehung $\varrho(\mathbf{P}_k^{-1}\mathbf{Q}_k) = \max\limits_{i=1,\ldots,n} \sqrt{[g_i(k)]} < 1$ sofort $g_i(k) < 1$ für alle $i = 1, \ldots, n$ folgt. Das tritt aber nach Hilfssatz 1, Behauptung c) genau dann ein, wenn k im Durchschnitt der Halbeneben (30) liegt, was wir beweisen wollten.

Durch der Verbindung der Sätze 8 und 9 bekommt man sofort einen Satz, der das den Optimalparameter k_0 enthaltende Gebiet weiter beschränkt.

10. *Es sei $\varrho(\mathbf{P}_1^{-1}\mathbf{Q}_1) < 1$. Der Parameter k_0, für den der Spektralradius $\varrho(\mathbf{P}_k^{-1}\mathbf{Q}_k)$ sein Minimum annimmt, liegt dann im Durchschnitt des Kreises $|k - 1| < 1$ und der Halbebenen (30).*

Ist die Matrix $\mathbf{P}_1^{-1}\mathbf{Q}_1$ singulär (das kommt bei gewissen Iterationverfahren, z.B. bei dem Gauss-Seidelschen Verfahren vor), d.h. wenn sie den Eigenwert $\lambda = 0$ hat, folgt aus dem vorhergehenden Satz sofort, dass der Optimalparameter k_0 im Durchschnitt des Kreises $|k - 1| < 1$ und der durch die Gleichung $-2k_1 + 1 < 1$ bestimmten Halbebene (das ist offensichtlich die durch die Gerade $k_1 = \frac{1}{2}$ begrenzte Halbebene, die den Koordinatenursprung nicht enthält) liegen muss. Dieses Gebiet ist in Abb. 5 mit der starken Linien gekennzeichnet.

Bei praktischen Berechnungen kennt man in der Regel die Eigenwerte der Matrix $\mathbf{P}_1^{-1}\mathbf{Q}_1$ nicht und man muss dann den Parameter k_0 mit Hilfe einer anderen Methode (z.B. probeweise) suchen (siehe auch [3]). Solange man wenigstens über die Lage irgendwelcher Eigenwerte der Matrix $\mathbf{P}_1^{-1}\mathbf{Q}_1$ eine Vorstellung hat, kann der Satz 10 manchmal das Gebiet, in dem wir den Optimalparameter k_0 suchen, verkleinern.

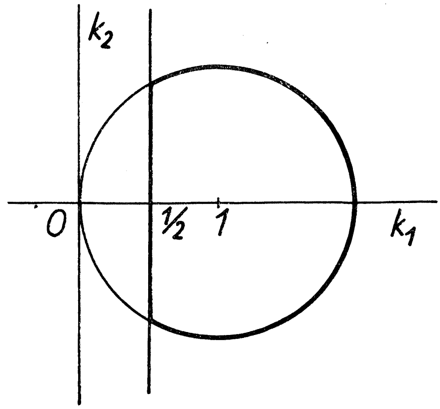

Abb. 5.

Zum Schluss beweisen wir noch den folgenden Satz:

11. *Enthält das Spektrum der Matrix $\mathbf{P}_1^{-1}\mathbf{Q}_1$ nur reelle und komplexe konjugierte Eigenwerte, dann ist der Optimalparameter k_0 eine reelle Zahl.*

Bemerkung 4. Wenn es sich z.B. um ein reelles Iterationsverfahren handelt, wo die Matrix $\mathbf{P}_1^{-1}\mathbf{Q}_1$ reell ist, sind die Voraussetzungen des Satzes 11 erfüllt und man kann also den Optimalparameter direkt im reellen Gebiete suchen (damit befasst sich die Arbeit [3]).

Beweis des Satzes 11. Enthält das Spektrum der Matrix nur reelle und komplexe konjugierte Eigenwerte, dann weiss man nach Satz 6, dass der Optimalparameter nur eine von den Zahlen k^{ij} oder k^{ijk} $(i, j = 1, \ldots, n)$ sein kann, solange Λ_i, Λ_j, Λ_k zu keiner Geraden gehören. Wir werden jetzt solche Punkte k^{ij}, k^{ijk} untersuchen, die nicht auf der reellen Achse liegen, d.h. $k_2^{ij} \neq 0$, $k_2^{ijk} \neq 0$, und beweisen, dass für diese Punkte der Ausdruck $\max_i \sqrt{[g_i(k)]}$ sein Minimum bezüglich k annehmen kann.

Zuerst beweist man, dass die Flächen $z = g_i(k)$, $z = g_j(k)$, wo $\Lambda_j = \bar{\Lambda}_i$ ist, symmetrisch bezüglich einer Ebene π sind, die die reelle Achse enthält und orthogonal zur komplexen Ebene ist. Es ist nämlich

$$g_i(k_1 + k_2 \mathrm{i}) = g_j(k_1 - k_2 \mathrm{i}),$$

wie man aus der Formel (3) leicht feststellen kann. Wenn ferner Λ_i reell ist, dann sind selbst die Punkte der Fläche $z = g_i(k)$ symmetrisch bezüglich der Ebene π, denn nach (3) gilt offensichtlich

$$g_i(k_1 + k_2\mathrm{i}) = g_i(k_1 - k_2\mathrm{i}) .$$

Nun werden wir einige Fälle unterscheiden.

I. Es sei k^{ij} ein Optimalparameter und k^{ij} sei keine reelle Zahl. Nach der Voraussetzung enthält das um 1 nach links verschobene Spektrum der Matrix $\boldsymbol{P}_1^{-1}\mathbf{Q}_1$ gemeinsam mit den Zahlen Λ_i, Λ_j auch die konjugierten Zahlen $\Lambda_{i'} = \bar{\Lambda}_i$, $\Lambda_{j'} = \bar{\Lambda}_j$.

a) Es sei $J_i > 0$, $J_j \geqq 0$. In Betracht auf die Beziehung (5) ist dann $k_2^{ij} < 0$ und es ist also auch $J_i k_2^{ij} < 0$. Diese Ungleichung ist aber offensichtlich mit der Ungleichung

$$(31) \qquad g_i(k^{ij}) < g_{i'}(k^{ij})$$

äquivalent (das folgt nämlich sofort aus (3) und aus dem Konjugiertsein der Zahlen Λ_i, $\Lambda_{i'}$). Der Punkt k^{ij} kann nicht der Optimalparameter sein, da sonst für jedes $l = 1, \ldots, n$ die Ungleichung

$$(32) \qquad g_i(k^{ij}) = g_j(k^{ij}) \geqq g_l(k^{ij})$$

gelten müsste, was ein Widerspruch mit (31) ist.

b) Es sei $J_i < 0$, $J_j \leqq 0$. Nach (5) ist $k_2^{ij} > 0$ und es ist also $J_i k_2^{ij} < 0$. Diese Ungleichung ist wieder mit der Ungleichung (31) äquivalent, was ein Widerspruch mit (32) ist.

c) Es sei $J_i > 0$, $J_j < 0$, $k_2^{ij} < 0$. Dann ist $J_i k_2^{ij} < 0$ und es gilt auch (31), was wieder ein Widerspruch mit (32) ist.

d) Es sei $J_i > 0$, $J_j < 0$, $k_2^{ij} > 0$. Dann ist $J_j k_2^{ij} < 0$, was eine mit der Ungleichung

$$g_j(k^{ij}) < g_{j'}(k^{ij})$$

äquivalente Ungleichung ist. Da ist wieder ein Widerspruch, denn wir haben vorausgesetzt, dass k^{ij} ein Optimalparameter ist.

Ein anderer Fall kann nicht eintretten. Dadurch ist bewiesen: wenn eine der Zahlen k^{ij} ein Optimalparameter ist, muss sie auf der reellen Achse liegen.

II. Man setze voraus, dass der Optimalparameter der Punkt $k_0 = k_1^0 + \mathrm{i}k_2^0$ ist, in dem sich die Flächen $g_{i_1}, g_{i_2}, \ldots, g_{i_r}$, $r \geqq 3$, schneiden, d.h. $g_{i_1}(k_0) = \ldots = g_{i_r}(k_0)$, und es sei $k_2^0 \neq 0$. Man bezeichne mit $\Lambda_{i_1'}, \Lambda_{i_2'}, \ldots, \Lambda_{i'_r}$, die mit den Zahlen Λ_{i_1}, $\Lambda_{i_2}, \ldots, \Lambda_{i_r}$ konjugierten Zahlen, d.h. $\Lambda_{i_1'} = \bar{\Lambda}_{i_1}, \ldots, \Lambda_{i_r'} = \bar{\Lambda}_{i_r}$. Nun können verschiedene Fälle eintreten:

a) Zwischen den Zahlen $\Lambda_{i_1}, \ldots, \Lambda_{i_r}$ existieren zwei Zahlen von verschiedenem

Vorzeichen, d.h. es sei z.B. $J_{i_1} > 0$, $J_{i_2} < 0$ und es sei $k_2^0 < 0$. Dann gilt $J_{i_1} k_2^0 < 0$, was mit der Ungleichung

$$g_{i_1}(k_0) < g_{i_1'}(k_0)$$

äquivalent ist, und das ist ein Widerspruch.

Wenn $k_2^0 > 0$ ist, ist $J_{i_2} k_2^0 < 0$ und es gilt also

$$g_{i_2}(k_0) < g_{i_2'}(k_0)\,,$$

was wieder ein Widerspruch ist.

b) Alle Zahlen $J_{i_1}, \ldots, J_{i_r}$ seien nichtnegativ. Dann muss wenigstens eine von diesen Zahlen positiv sein, denn anderenfalls würden alle Punkte $A_{i_1}, \ldots, A_{i_r}$ auf der gleichen Geraden liegen und es würde kein Schnittpunkt k_0 der Geraden $p_{i_j i_k}$, $j = 1, \ldots, r$, $k = 1, \ldots, r$ existieren. Es sei also z.B.

$$J_{i_1} > 0,\ J_{i_2} \geqq 0, \ldots, J_{i_r} \geqq 0\,.$$

Wenn nun $k_2^0 < 0$ ist, ist $J_{i_1} k_2^0 < 0$ und es gilt also

$$g_{i_1}(k_0) < g_{i_1'}(k_0)\,,$$

was ein Widerspruch ist.

Es sei $k_2^0 > 0$. Nach der Formel (5) gilt $k_2^{i_j i_k} \leqq 0$ für alle $j, k = 1, \ldots, r$, sodass der Punkt k_0 in keinem von den Dreiecken mit den Eckpunkten $k^{i_j i_k}$ liegen kann. Er kann also auch nicht im kleinsten konvexen die Punkte $k^{i_j i_k}$ enthaltenden Vielecke liegen. Das ist aber ein Widerspruch mit der Behauptung des Satzes 7. Dadurch ist der Satz 11 bewiesen.

Zur Illustration der oben abgeleiteten Ergebnisse führen wir nun ein numerisches Beispiel an: Man setzt voraus, dass die Matrix $\mathbf{P}_1^{-1}\mathbf{Q}_1$ die folgenden Eigenwerte hat:

$$\lambda_1 = 0{,}4 + 0{,}2\mathrm{i}\,, \quad \lambda_2 = 0{,}2 + 0{,}6\mathrm{i}\,, \quad \lambda_3 = 0{,}6 + 0{,}7\mathrm{i}\,.$$

Offensichtlich ist $|\lambda_1| < |\lambda_2| < |\lambda_3|$. Der Spektralradius ist dann gleich $|\lambda_3| = 0{,}92$, was eine sehr schlechte Konvergenz des Iterationsverfahrens (1) für $k = 1$ darstellt. Wenn man die Ergebnise der Arbeit [3] anwendet, sieht man, dass die Voraussetzungen des Satzes 2.7 aus [3] erfüllt sind, und daraus folgt, dass der reelle Optimalparameter k_0 der Zahl

$$k_0 = k_3 = \frac{|A_3|^2}{|\mathrm{Re}\, A_3|} = 1{,}62$$

gleich ist und dass für den entsprechenden Spektralradius

$$\varrho(\mathbf{P}_{k_0}^{-1}\mathbf{Q}_{k_0}) = 0{,}87$$

gilt. Das heisst, dass die Methode (1) für $k = 1{,}62$ etwas schneller konvergieren wird, als die ursprüngliche Methode für $k = 1$.

Wenn man jetzt den Optimalparameter im komplexen Gebiete sucht, dann stellt man nach der Berechnung der Zahlen k^{12}, k^{13}, k^{23}, k^{123} durch dem Satz 6 dieser Arbeit fest, dass der Optimalparameter die Zahl

$$k_0 = k^{13} = 0{,}6 - 0{,}49\mathrm{i}$$

ist. Der entsprechende Spektralradius ist dann die Zahl

$$\varrho(\mathbf{P}_{k_0}^{-1}\mathbf{Q}_{k_0}) = 0{,}375\,,$$

was schon eine sehr wesentliche Konvergenzbeschleunigung darstellt.

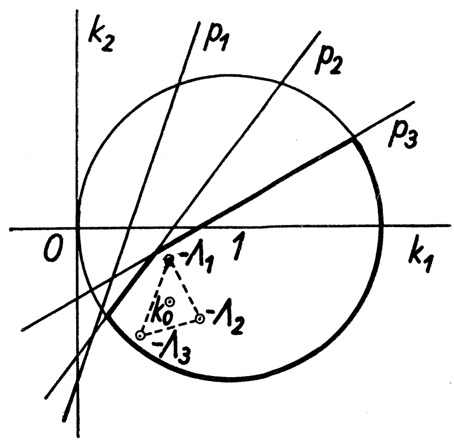

Abb. 6.

In Abb. 6 ist das Gebiet aufgezeichnet, in dem nach dem Satz 9 der Optimalparameter k_0 liegt. Der durch die Geraden

$$\begin{aligned} p_1 &\equiv -1{,}2k_1 + 0{,}4k_2 + 0{,}4 = 0\,, \\ p_2 &\equiv -1{,}6k_1 + 1{,}2k_2 + 1 = 0\,, \\ p_3 &\equiv -0{,}8k_1 + 1{,}4k_2 + 0{,}65 = 0 \end{aligned}$$

begrenzte Durchschnitt der Halbebenen (25), die den Koordinatenursprung nicht enthalten, und des Kreises $|k - 1| < 1$ ist durch die starke Linie begrenzt.

Bemerkung 5. Bei der praktischen Anwendung der Methode kann die folgende angenäherte Behauptung nützlich sein: Wenn die Eigenwerte λ_i der Matrix $\mathbf{P}_1^{-1}\mathbf{Q}_1$ in irgendeinem verhältnismässig kleinen Gebiete des Einheitskreises konzentriert sind, d.h. wenn auch die Zahlen Λ_i in einem kleinen Gebiet liegen, dann liegt der Optimalparameter k_0 in der Nähe der Zahlen $-\Lambda_i$. Dies ist aus dem Hilfssatz 1, Behauptung d), sichtbar, da die Funktionen g_i in den Punkten $-\Lambda_i$ ihr Minimum (das gleich Null ist) annehmen. Eine solche Situation tritt z.B. in dem oben angeführten Beispiele ein. In Abb. 6 sieht man, dass der Optimalparameter direkt im Dreiecke mit den Ecken $-\Lambda_1$, $-\Lambda_2$, $-\Lambda_3$ liegt.

Literaturverzeichnis

[1] *Isaacson, E., Keller, H. B.:* Analysis of Numerical Methods, John Wiley & Sons, Inc.; New York, London, Sydney, 1966.

[2] *Šisler, M.:* Über die Konvergenzbeschleunigung verschiedener Iterationsverfahren. Aplikace matematiky 12, 255—267 (1967).

[3] *Šisler, M.:* Über eine Relaxationsmethode. Aplikace matematiky 13, 478—488 (1968).

Souhrn

O URYCHLOVÁNÍ KONVERGENCE KOMPLEXNÍCH ITERAČNÍCH PROCESŮ

MIROSLAV ŠISLER

Práce se zabývá urychlováním konvergence iteračních procesů pro řešení soustavy m lineárních rovnic $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$. Předpokládá se, že je dán iterační proces definovaný vzorcem

$$\boldsymbol{x}_{\nu+1} = \boldsymbol{P}_k^{-1}\boldsymbol{Q}_k\boldsymbol{x}_\nu + \boldsymbol{P}_k^{-1}\boldsymbol{b}\,, \quad \nu = 0, 1, 2, \ldots,$$

kde $\boldsymbol{P}_k = k\boldsymbol{P}_1$, $\boldsymbol{Q}_k = (k - 1)\,\boldsymbol{P}_1 + \boldsymbol{Q}_1$, $k \neq 0$. Přitom je $\boldsymbol{A} = \boldsymbol{P}_1 - \boldsymbol{Q}_1$ jistý rozklad matice $\boldsymbol{A}$ takový, že spektrální poloměr matice $\boldsymbol{P}_1^{-1}\boldsymbol{Q}_1$ je menší než 1 a k je komplexní parametr. (Pro $k = 1$ dostáváme zřejmě iterační proces odpovídající rozkladu $\boldsymbol{A} = \boldsymbol{P}_1 - \boldsymbol{Q}_1$.)

V článku je řešena otázka volby optimálního parametru k v tom smyslu, aby byl spektrální poloměr matice $\boldsymbol{P}_k^{-1}\boldsymbol{Q}_k$ minimální, tj. aby zkoumaný iterační proces konvergoval co nejrychleji. V článku je dokázáno, že optimální parametr k leží vždy v kruhu komplexní roviny o poloměru rovném 1 a středu v bodě 1, přičemž je vždy roven některému z čísel k^{ij}, k^{ijk}, $i, j, k = 1, \ldots, n$, jejichž reálné a imaginární části jsou dány vzorci (4), (5) a (21') (zde značí $\Lambda_i = R_i + \mathrm{i}J_i$, $\Lambda_i = \lambda_i - 1$, $i = 1, \ldots, n$, přičemž λ_i jsou vlastní čísla matice $\boldsymbol{P}_1^{-1}\boldsymbol{Q}_1$).

Dále je dokázáno, že obsahuje-li spektrum matice $\boldsymbol{P}_1^{-1}\boldsymbol{Q}_1$ jen reálná a komplexně sdružená vlastní čísla (tj. např. v případě reálného iteračního procesu), je optimální parametr reálný, tj. leží v intervalu $(0,2)$.

V článku je uveden jeden numerický příklad.

Anschrift des Verfassers: RNDr. *Miloslav Šisler*, CSc., Matematický ústav ČSAV v Praze, Praha 1, Žitná 25.