Zeitschrift: Commentarii Mathematici Helvetici

Herausgeber: Schweizerische Mathematische Gesellschaft

Band: 29 (1955)

Artikel: Relaxationsmethoden bester Strategie zur Lösung linearer Gleichungssysteme.

Autor: Stiefel, E.

DOI: https://doi.org/10.5169/seals-23284

Nutzungsbedingungen

Die ETH-Bibliothek ist die Anbieterin der digitalisierten Zeitschriften auf E-Periodica. Sie besitzt keine Urheberrechte an den Zeitschriften und ist nicht verantwortlich für deren Inhalte. Die Rechte liegen in der Regel bei den Herausgebern beziehungsweise den externen Rechteinhabern. Das Veröffentlichen von Bildern in Print- und Online-Publikationen sowie auf Social Media-Kanälen oder Webseiten ist nur mit vorheriger Genehmigung der Rechteinhaber erlaubt. Mehr erfahren

Conditions d'utilisation

L'ETH Library est le fournisseur des revues numérisées. Elle ne détient aucun droit d'auteur sur les revues et n'est pas responsable de leur contenu. En règle générale, les droits sont détenus par les éditeurs ou les détenteurs de droits externes. La reproduction d'images dans des publications imprimées ou en ligne ainsi que sur des canaux de médias sociaux ou des sites web n'est autorisée qu'avec l'accord préalable des détenteurs des droits. En savoir plus

Terms of use

The ETH Library is the provider of the digitised journals. It does not own any copyrights to the journals and is not responsible for their content. The rights usually lie with the publishers or the external rights holders. Publishing images in print and online publications, as well as on social media channels or websites, is only permitted with the prior consent of the rights holders. Find out more

Download PDF: 05.10.2025

ETH-Bibliothek Zürich, E-Periodica, https://www.e-periodica.ch

Relaxationsmethoden bester Strategie zur Lösung linearer Gleichungssysteme

von E. Stiefel, Zürich

Herrn Prof. Dr. M. Plancherel zum 70. Geburtstag in Dankbarkeit gewidmet

I. Die Auflösung eines gegebenen Gleichungssystems

$$A x = k \tag{1}$$

mit symmetrischer und positiv definiter Matrix A ist äquivalent mit der Aufgabe, das Minimum der quadratischen Funktion

$$F(x) = \tfrac{1}{2}(x, A x) - (k, x) \tag{2}$$

zu suchen. Hierin bedeutet das Komma die Bildung des skalaren Produkts der beiden in der Klammer stehenden Vektoren. Im euklidischen Raum R der Vektoren x sind die Niveauflächen dieser Funktion ähnliche und ähnlich gelegene Ellipsoide, deren gemeinsames Zentrum der Lösungspunkt $A^{-1}k$ des Gleichungssystems (1) ist.

Zur Lösung der gestellten Aufgabe wird das *Verfahren des stärksten Abstiegs* (Gradientenmethode) empfohlen. Man konstruiert ausgehend von einem Näherungspunkt x_0 die von x_0 auslaufende orthogonale Trajektorie der Niveauflächen; diese Kurve endigt im gesuchten Lösungspunkt. Die Durchführung dieses Ansatzes führt auf folgende Formeln. Der Gradient von $F(x)$ ist gegeben durch

$$\operatorname{grad} F = A x - k \ , \tag{3}$$

also hat man das System von linearen Differentialgleichungen zu lösen:

$$\frac{dx}{dt} = k - A x \ , \tag{4}$$

wobei t ein Parameter auf der Trajektorie ist. Die rechte Seite von (4) hat eine einfache Bedeutung. Sie gibt das Residuum an, das übrigbleibt, wenn man x als Versuchspunkt in die Gleichungen (1) einsetzt. Wir nennen daher

$$r = k - A x \tag{5}$$

den zu x gehörigen Residuenvektor und haben

$$\frac{dx}{dt} = r \,. \tag{6}$$

Dieses System von Differentialgleichungen kann nach einer der klassischen numerischen Methoden gelöst werden. Die einfachste Möglichkeit liefert das Eulersche Polygonzugverfahren. Man wählt eine Einteilung t_i der t-Achse und rechnet nach der Vorschrift

$$\Delta x_i = r_i \Delta t_i \qquad (i = 0, 1, 2, \cdots) \,, \tag{7}$$

wobei

$$\Delta t_i = t_{i+1} - t_i \qquad \Delta x_i = x_{i+1} - x_i$$

gesetzt wurde. Es ist für das folgende bequemer, den reziproken Wert q_i von Δt_i einzuführen:

$$\Delta x_i = \frac{1}{q_i} r_i \,, \qquad q_i > 0 \,. \tag{8}$$

Die verschiedenen Gradientenverfahren unterscheiden sich nun durch die Wahl der q_i, und diese Wahl ist von ausschlaggebender Bedeutung. Teilt man nämlich fein ein, so erhält man zwar einen Polygonzug, der nahe der gewünschten Trajektorie verläuft, aber erst nach sehr vielen Schritten in die Nähe des Lösungspunktes führt; bei allzu grober Einteilung kann der Polygonzug divergieren. Am bekanntesten sind folgende beiden speziellen Gradientenverfahren.

1. Wählt man eine gleichmäßige Einteilung der t-Achse, das heißt, nimmt man in (8) für q_i eine feste von i unabhängige Zahl q, so erhält man einen Algorithmus, der in der deutschen Literatur als *Gesamtschrittverfahren* bezeichnet wird.

2. Zur Erklärung des sogenannten *optimalen* Gradientenverfahrens brauchen wir den *Fehlervektor*

$$y = A^{-1}k - x \,, \tag{9}$$

der zu einem Näherungspunkt x gehört, und benötigen ein Maß für die Größe dieses Fehlers. Als solches werden wir wahlweise benutzen

a) Das euklidische Maß

$$\varphi(y) = (y, y) \,. \tag{10}$$

b) Das durch die quadratische Form in (2) gegebene „*A-Maß*“

$$\varphi'(y) = (y, Ay) \,. \tag{11}$$

Wegen $r = Ay$ gilt noch

$$\varphi'(y) = (y, r) = (r, A^{-1}r) \ . \tag{12}$$

c) Das Residuenmaß

$$\varphi''(y) = (r, r) = (Ay, Ay) \ . \tag{13}$$

Nur dieses letzte Maß kann bei gegebenem Näherungspunkt x numerisch berechnet werden. Alle Maße sind positiv definit, das heißt aus $\varphi(y) = 0$ folgt $y = 0$.

Das optimale Gradientenverfahren beruht nun auf folgender Grundidee der Relaxationsrechnung. Wir betrachten in (8) den i-ten Schritt, also

$$x_{i+1} = x_i + \frac{1}{q_i} r_i \ ,$$

was geometrisch ein geradliniges Fortschreiten von x_i aus in der Richtung r_i bedeutet. Man wähle nun q_i und damit x_{i+1} auf diesem geradlinigen Weg so, daß der zugehörige Fehler y_{i+1} möglichst klein wird. Bei Zugrundelegung des A-Maßes ergibt dies:

$$y_{i+1} = y_i - \frac{1}{q_i} r_i$$

$$\varphi'(y_{i+1}) = \varphi'(y_i) - \frac{2}{q_i}(r_i, r_i) + \frac{1}{q_i^2}(r_i, Ar_i) \ ,$$

und somit

$$q_i = \frac{(r_i, Ar_i)}{(r_i, r_i)} \ . \tag{14}$$

q_i muß also gleich dem Rayleighschen Quotient des Residuenvektors gewählt werden. Verwendet man jedoch das Residuenmaß, so ergibt sich

$$q_i = \frac{(Ar_i, Ar_i)}{(r_i, Ar_i)} \ , \tag{15}$$

was ein Schwarzscher Quotient ist. Die analoge Formel für das euklidische Maß

$$q_i = \frac{(r_i, r_i)}{(r_i, A^{-1}r_i)} \tag{16}$$

ist nicht brauchbar, da A^{-1} unbekannt ist. Bezeichnet man die drei Werte (14) (15) (16) beziehlich mit q_i', q_i'', q_i, so gilt wegen der bekannten

Ungleichungen zwischen Schwarzschen Quotienten

$$q_i \leqslant q_i' \leqslant q_i'' \,. \tag{17}$$

Die drei zugehörigen Näherungspunkte x_{i+1}'', x_{i+1}', x_{i+1} liegen also von x_i aus in der Richtung r_i gesehen in *dieser* Reihenfolge. Da x_{i+1} der Fußpunkt des Lotes ist, das vom Lösungspunkt $A^{-1}k$ auf unseren geradlinigen Weg gefällt werden kann, ergibt sich, daß x_{i+1}' und x_{i+1}'' im Sinne der euklidischen Metrik näher am Lösungspunkt liegen als x_i. Bei beiden Varianten (14) und (15) des optimalen Gradientenverfahrens nähert man sich also dem Lösungspunkt in monotoner Weise. Trotzdem beschreibt man einen Zickzack-Kurs, was daraus hervorgeht, daß zum Beispiel bei Verwendung von (14) der Eulersche Polygonzug lauter rechte Winkel hat.

Von *M. Hestenes* [1] (vgl. das Literaturverzeichnis am Schluß) ist daher vorgeschlagen worden, durch Hinzufügen eines willkürlichen Koeffizienten β_i mit $0<\beta_i<1$ in der Formel (8) den Polygonzug zu glätten. Man rechnet dann also etwa nach der Vorschrift

$$\Delta x_i = \beta_i \frac{1}{q_i} r_i \,,$$

wobei q_i durch (14) gegeben ist. Numerische Beispiele wurden von *M. L. Stein* [2] gerechnet.

Im Sinn der Theorie der Spiele liefern die optimalen Gradientenmethoden eine *Taktik* zur Auflösung der gegebenen linearen Gleichungen, indem bei jedem einzelnen Schritt (ohne Rücksicht auf Vorgeschichte und Weiterentwicklung des Iterationsprozesses) das Beste getan wird. Im folgenden soll die zugehörige *Strategie* entwickelt werden, die also den Prozeß auf längere Sicht hin möglichst günstig steuert. Wir betrachten alle Gradientenprozesse (8), die eine vorgegebene Anzahl n von Schritten aufweisen. Jeder von ihnen ist charakterisiert durch eine Folge von positiven Koeffizienten

$$q_0, q_1, \ldots, q_{n-1} \tag{18}$$

und führt zu einem Endpunkt x_n mit zugehörigem Fehlervektor y_n. Es soll nun die Koeffizientenfolge (18) so bestimmt werden, daß y_n möglichst klein wird im Sinne eines unserer Fehlermaße. Den Endpunkt x_n nennen wir einen *Punkt bester Strategie.*

II. Um unser Strategieproblem zu lösen, wollen wir zunächst die iterative Rechenvorschrift (8) in einen geschlossenen Ausdruck für den Endpunkt x_n verwandeln. Ohne Beeinträchtigung der Allgemeinheit können

wir annehmen, daß der Anfangspunkt x_0 der Nullpunkt ist. Nehmen wir dann noch das Polynom n-ten Grades in einer Unbestimmten λ

$$R_n(\lambda) = \left(1 - \frac{\lambda}{q_0}\right)\left(1 - \frac{\lambda}{q_1}\right) \cdots \left(1 - \frac{\lambda}{q_{n-1}}\right) \tag{19}$$

zu Hilfe, so ergibt sich leicht durch vollständige Induktion

$$x_n = \frac{1 - R_n(A)}{A} k \tag{20}$$

mit dem zugehörigen Residuum

$$r_n = R_n(A) k \ . \tag{21}$$

Das Polynom $R_n(\lambda)$ ist durch folgende beiden Eigenschaften charakterisiert. Erstens nimmt es für $\lambda = 0$ den Wert Eins an

$$R_n(0) = 1 \ , \tag{22}$$

und zweitens besitzt es die Zahlen unserer Folge (18) als Nullstellen. Umgekehrt ist durch die Wahl eines Polynoms $R_n(\lambda)$ mit lauter *positiven* Nullstellen eine Folge (18) und damit ein Gradientenprozeß gegeben.

Überdies aber gibt das Polynom $R_n(\lambda)$ Auskunft über die bei x_n erreichte Genauigkeit der Gleichungsauflösung. Rechnen wir nämlich im Koordinatensystem der Hauptachsen von A und bezeichnen mit λ_j die Eigenwerte von A sowie beziehlich mit r_{nj}, k_j die j-ten Komponenten von r_n, k, so folgt aus (21)

$$r_{nj} = R_n(\lambda_j) k_j \ . \tag{23}$$

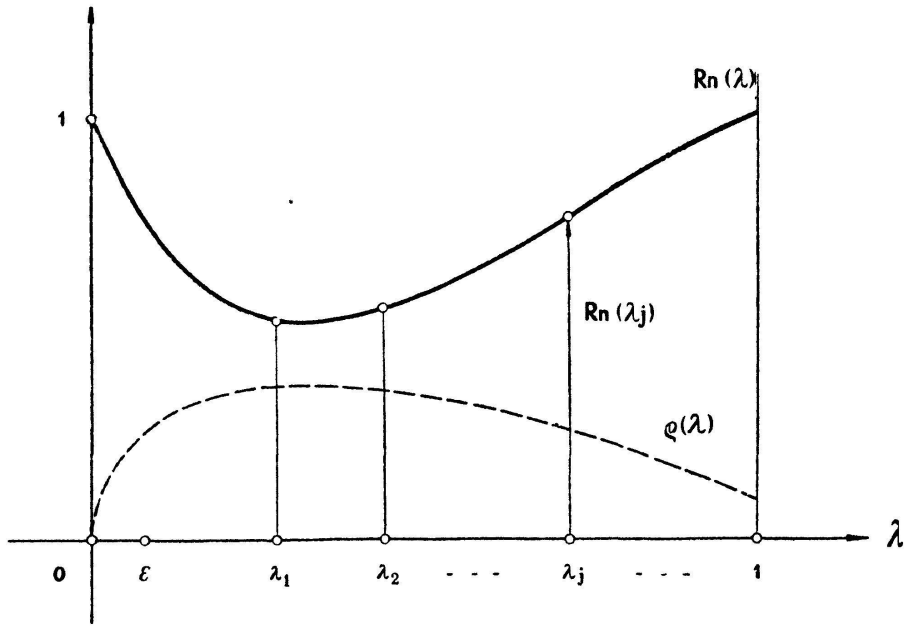

Fig. 1

Wir zeichnen nun über einer λ-Achse eine graphische Darstellung des Polynoms $R_n(\lambda)$ und markieren auf der λ-Achse die Eigenwerte. Dann geben die zugehörigen Ordinaten an, um wieviel Prozent die betreffende Komponente des anfänglichen Residuenvektors k beim Gradientenprozeß reduziert wurde. Wegen (22) werden dabei Eigenwerte, die nahe am Nullpunkt liegen, Schwierigkeiten machen. Dies ist zu erwarten, da in einem solchen Fall das gegebene Gleichungssystem schlecht konditioniert ist.

Bereits diese mehr qualitativen Überlegungen geben einen Hinweis, wie die Koeffizienten q_i für einen Gradientenprozeß guter Strategie zu wählen sind. Sie müssen ordentlich verteilt in dasjenige Intervall der λ-Achse gelegt werden, in welchem man Eigenwerte von A erwartet, damit $R_n(\lambda)$ in diesem Intervall kleine Werte annimmt. Für gut konditionierte Systeme (1) wird diese Regel oft genügen, da bei solchen leicht Schranken für die Eigenwerte a priori angegeben werden können. Beim *Gesamtschrittverfahren* von Nr. I hat das Polynom $R_n(\lambda)$ eine einzige n-fache Nullstelle für $\lambda = q$; in ihrer Umgebung werden die Residuen rasch eliminiert, während an anderen Stellen der λ-Achse das Verfahren nicht sehr wirkungsvoll ist.

Wegen (21) und (23) bezeichnen wir $R_n(\lambda)$ hinfort als *Residuenpolynom.*

III. Die exakte Lösung des Strategieproblems wird sehr erleichtert, wenn wir die drei Fehlermaße von Nr. I als Spezialfälle eines allgemeineren Maßes ψ darstellen. Um ein solches zu definieren, brauchen wir eine obere Schranke für die Eigenwerte von A, die nicht sehr genau zu sein braucht und die wir unbeschadet der Allgemeinheit zu Eins normieren. Ferner wählen wir im Intervall $0 < \lambda < 1$ eine beliebige positive *Dichtefunktion* $\varrho(\lambda)$ so, daß das Integral

$$\int_0^1 \frac{\varrho(\lambda)}{\lambda} d\lambda$$

existiert. Dann definieren wir

$$\psi_n = \int_0^1 \frac{R_n(\lambda)^2}{\lambda} \varrho(\lambda) d\lambda \ . \tag{24}$$

Unser altes Maß φ' zum Beispiel entspricht der Wahl

$$\varrho(\lambda) = \sum_{(j)} k_j^2 \delta(\lambda - \lambda_j) \ , \tag{25}$$

wobei δ die Diracsche δ-Funktion bedeutet.

In der Tat ist dann

$$\psi_n = \sum_{(j)} k_j^2 \int_0^1 \frac{R_n(\lambda)^2}{\lambda} \delta(\lambda - \lambda_j) d\lambda = \sum_{(j)} k_j^2 \frac{R_n(\lambda_j)^2}{\lambda_j}$$

$$= \sum_{(j)} r_{nj} \frac{r_{nj}}{\lambda_j} = (r_n, A^{-1} r_n) = \varphi'(y_n) \ ,$$

wobei noch (23) und (12) benutzt wurden. Analog entsprechen die Maße φ, φ'' den Dichten

$$\frac{1}{\lambda} \Sigma k_j^2 \delta(\lambda - \lambda_j) \quad \text{und} \quad \lambda \Sigma k_j^2 \delta(\lambda - \lambda_j) \ . \tag{26}$$

In allen drei Fällen macht die Dichtefunktion an den Eigenwerten δ-Sprünge; wir werden im folgenden aber durchaus auch stetige Dichtefunktionen zulassen. In einem solchen Fall sprechen wir von einem *kontinuierlichen* Fehlermaß, während die φ-Maße von Nr. I als *diskontinuierlich* bezeichnet werden sollen. Das am Ende von Nr. I gestellte Strategieproblem nimmt nun folgende einfache Form an:

Unter allen Polynomen n-ten Grades $R_n(\lambda)$ mit positiven Nullstellen, welche der Nebenbedingung $R_n(0) = 1$ genügen, ist dasjenige zu finden, welches das Integral (24) bei gegebener Dichtefunktion $\varrho(\lambda)$ minimal macht. Zur Lösung dieser Extremalaufgabe konstruieren wir die Folge $P_\nu(\lambda)$ von Orthogonalpolynomen zur Dichtefunktion $\varrho(\lambda)$:

$$\int_0^1 P_\mu(\lambda) P_\nu(\lambda) \varrho(\lambda) d\lambda = 0 \quad \text{für} \quad \mu \neq \nu \ . \tag{27}$$

P_ν habe den Grad ν und werde durch die Bedingung

$$P_\nu(0) = 1 \tag{28}$$

eindeutig festgelegt. (Diese Normierungsbedingung ist also verschieden von der üblichen, bei welcher verlangt wird, daß das Normierungsintegral eins sei.) Dann liefert $R_n(\lambda) = P_n(\lambda)$ die Lösung des Extremalproblems.

Beweis. Zunächst erfüllt $P_n(\lambda)$ unsere Konkurrenzbedingungen, denn die Nullstellen $\pi_0, \pi_1, \ldots, \pi_{n-1}$ dieses Polynoms sind nach bekannten Sätzen über Orthogonalpolynome alle reell und liegen im Intervall $0 < \lambda < 1$. Außerdem ist wegen (28) die Nebenbedingung erfüllt. Weiter verwenden wir das Hilfsmittel der allgemeinen Gaußschen Quadratur (vgl. Szegö [3], S. 46). Es existieren positive Integrationsgewichte $p_0, p_1, \ldots, p_{n-1}$, so daß für jedes Polynom $P(\lambda)$ vom Höchstgrad $(2n - 1)$ gilt

$$\int_0^1 P(\lambda) \varrho(\lambda) d\lambda = \sum_{j=0}^{n-1} p_j P(\pi_j) \ .$$

Ist nun $R_n(\lambda)$ ein den Konkurrenzbedingungen genügendes Polynom, so wenden wir diese Formel auf

$$P(\lambda) = \frac{R_n(\lambda)^2 - 1}{\lambda}$$

an, was in der Tat ein Polynom vom Grad $(2n - 1)$ ist. Es folgt

$$\psi_n = \int_0^1 \frac{R_n(\lambda)^2}{\lambda} \varrho(\lambda) d\lambda = \int_0^1 \frac{\varrho(\lambda)}{\lambda} d\lambda + \sum_{j=0}^{n-1} p_j \frac{R_n(\pi_j)^2 - 1}{\pi_j} .$$

Die rechte Seite wird minimal, wenn $R_n(\pi_j) = 0$ ist, das heißt, wenn $R_n(\lambda) = P_n(\lambda)$ ist. Für den Wert des Minimums ergibt sich noch

$$\psi_n = \int_0^1 \frac{\varrho(\lambda)}{\lambda} d\lambda - \sum_{j=0}^{n-1} \frac{p_j}{q_j} , \tag{29}$$

falls wir wie früher die Nullstellen von R_n wieder mit $q_0, q_1, \ldots, q_{n-1}$ bezeichnen. Damit ist der Beweis beendigt. Übrigens ist die rechte Seite von (29) einfach der Fehler, den man begeht, wenn man das Integral der Funktion $1/\lambda$ mit Hilfe der Gaußschen Quadratur angenähert berechnet. Wir formulieren als vorläufiges Resultat

Satz 1. *Zu jeder Dichtefunktion $\varrho(\lambda)$ gehört ein Gradientenverfahren bester Strategie, welches eine gegebene Anzahl n von Schritten aufweist. Es wird gefunden, indem man die zur Dichtefunktion $\varrho(\lambda)$ gehörige Schar orthogonaler Polynome bis zu einem Polynom n-ten Grades aufbaut und für die Koeffizienten q_i in der Rechenvorschrift* (8) *die Nullstellen dieses Polynoms in beliebiger Reihenfolge nimmt.*

Ein Beispiel möge dies erläutern. Wir wählen als Dichte

$$\varrho(\lambda) = \sqrt{\lambda(1 - \lambda)} .$$

Führt man die Hilfsvariable φ durch

$$\cos \varphi = 1 - 2\lambda$$

ein und bildet die Funktionen

$$R_n(\lambda) = \frac{1}{n+1} \frac{\sin (n+1)\varphi}{\sin \varphi} , \tag{30}$$

so bestätigt man leicht, daß sie orthogonal sind in bezug auf die gegebene Dichte und daß $R_n(\lambda)$ ein Polynom vom Grad n ist, welches für $\lambda = 0$ den Wert Eins annimmt. Die Nullstellen von $R_n(\lambda)$ können konstruiert werden, indem man einen Kreis über der Strecke von 0 bis 1 der λ-Achse als Durchmesser zeichnet und ihm ein reguläres $2(n + 1)$-Eck einbe-

schreibt, das eine Ecke im Nullpunkt hat. Die Nullstellen — (und damit die q-Werte in (8)) — sind dann die Abszissen der Eckpunkte (mit Ausnahme von 0 und 1). Die Polynome (30) sind Tschebyscheffsche Sinuspolynome; sie sind tabelliert bei *Lanzcos* [4]. In der dortigen Bezeichnung ist

$$R_n(\lambda) = \frac{1}{n+1} S_n(2 - 4\lambda) \ .$$

Für unser Fehlermaß (29) ergibt sich noch

$$\psi_n = \int_0^1 R_n(\lambda)^2 \sqrt{\frac{1-\lambda}{\lambda}}\, d\lambda = \frac{1}{n+1} \frac{\pi}{2} \ .$$

Wie ersichtlich berücksichtigt dieses Maß die zu kleinen Eigenwerten gehörigen Residuen mit einem großen Gewicht, während die zu höheren Eigenwerten gehörigen kaum etwas beitragen. Diese Art der Fehlermessung ist vorteilhaft, wenn man Randwertaufgaben der mathematischen Physik löst, indem man ein Punktgitter über das Grundgebiet legt und Differenzenrechnung anwendet. Es entsteht dann ein schlecht konditioniertes Gleichungssystem, das also Eigenwerte in der Nähe des Nullpunkts hat. Überdies aber sind meistens die zu diesen kleinen Eigenwerten gehörigen Residuenkomponenten groß, so daß sie eben beim Relaxationsprozeß mit großem Gewicht zu berücksichtigen sind. Diese Erscheinung wurde in einer früheren Arbeit [5] als „Käfig" bezeichnet und ausführlich diskutiert.

Ähnliche Polynome wie (30) benützt *C. Lanczos* [6] bei einer anderen Methode zur Gleichungsauflösung.

Bei *gut konditionierten* Systemen wird eine untere Schranke $\varepsilon > 0$ für die Eigenwerte zum voraus bekannt sein; man wird dann $\varrho(\lambda) = 0$ wählen im Intervall $0 < \lambda < \varepsilon$, weil dort keine Residuen liquidiert werden müssen. Das obige Beispiel wäre also dann so zu modifizieren, daß man den Kreis über der Strecke von ε bis 1 konstruiert und das reguläre Polygon einbeschreibt. Man erhält so ein Relaxationsverfahren, das im wesentlichen mit einer von *D. Young* [7] entwickelten Methode übereinstimmt, die auf Untersuchungen von *Shortley* und *Flanders* beruht. Die Youngsche Methode versagt jedoch für $\varepsilon = 0$ und ist daher nicht sehr wirksam für kleine ε.

IV. Das in Satz 1 niedergelegte Resultat ist noch in mehrfacher Hinsicht unvollständig. Erstens ist der im Punkt x_n bester Strategie endigende Polygonzug nicht eindeutig bestimmt, sondern abhängig von der gewählten Reihenfolge der Nullstellen des Residuenpolynoms $R_n(\lambda)$.

Die Zwischenpunkte x_i des Polygonzugs sind durchaus keine Punkte bester Strategie und daher als Näherungslösung des Gleichungssystems meistens unbrauchbar. Mit anderen Worten, man darf den Gradientenprozeß nicht abbrechen, bevor der n-te Schritt wirklich ausgeführt ist.

Zweitens ergibt Satz 1 nichts im Fall der Dichten (25) (26), die zu den diskontinuierlichen Fehlermaßen φ, φ', φ'' gehören und uns eigentlich am meisten interessieren. In der Tat sind ja diese Dichten mit Hilfe der Eigenwerte der Gleichungsmatrix gebildet und diese Eigenwerte können nicht als bekannt angesehen werden.

Alle diese Nachteile werden behoben, wenn wir den Gradientenprozeß etwas modifizieren, indem wir statt der Nullstellen der Orthogonalpolynome die Rekursionsformel benutzen, welche zwischen ihnen besteht. Es seien

$$R_0(\lambda), R_1(\lambda), \ldots, R_i(\lambda), \ldots, R_n(\lambda) \tag{31}$$

die Orthogonalpolynome zur vorgegebenen Dichte $\varrho(\lambda)$, welche die Normierungsbedingung

$$R_i(0) = 1 \tag{32}$$

erfüllen. Zwischen je dreien besteht eine Rekursionsformel (vgl. [3], S. 41), die man erhält, indem man $\lambda R_i(\lambda)$ nach den Orthogonalpolynomen entwickelt:

$$\lambda R_i(\lambda) = -q_i R_{i+1}(\lambda) + t_i R_i(\lambda) - p_i R_{i-1}(\lambda) \ . \tag{33}$$

q_i, t_i, p_i sind die Entwicklungskoeffizienten und haben nichts zu tun mit Größen, die früher mit denselben Buchstaben bezeichnet wurden. Wegen (32) folgt

$$t_i = p_i + q_i \ ,$$

also

$$\lambda R_i(\lambda) = -q_i R_{i+1}(\lambda) + (p_i + q_i) R_i(\lambda) - p_i R_{i-1}(\lambda) \ , \tag{34}$$

wobei noch

$$R_0 = 1 \ , \quad p_0 = 0 \tag{35}$$

zu setzen ist.

Wir konstruieren nun eine neue Folge von Näherungspunkten x_i nach der Vorschrift

$$x_0 = 0 \ , \quad x_{i+1} = x_i + \Delta x_i \ , \quad \Delta x_i = \frac{1}{q_i}(r_i + p_i \Delta x_{i-1}) \ , \tag{36}$$

wobei r_i wieder der zu x_i gehörige Residuenvektor ist. Wir behaupten, daß dieser Polygonzug in demselben Punkt x_n endigt, wie der nach Satz 1 mit Hilfe des Polynoms $R_n(\lambda)$ konstruierte Gradientenprozeß.

Beweis. Wir zeigen durch vollständige Induktion

$$r_{i+1} = R_{i+1}(A)k \ . \tag{37}$$

Zunächst folgt aus

$$r_{i+1} = k - A x_{i+1} \ ,$$
$$r_i \ = k - A x_i \ ,$$

durch Subtraktion

$$r_{i+1} - r_i = - A \Delta x_i \ . \tag{38}$$

Somit

$$r_{i+1} = r_i - A \Delta x_i = r_i - \frac{1}{q_i} A r_i - \frac{p_i}{q_i} A \Delta x_{i-1} \ .$$

Indem man (38) noch einmal benutzt

$$r_{i+1} = r_i - \frac{1}{q_i} A r_i + \frac{p_i}{q_i} (r_i - r_{i-1}) \ .$$

Nach Induktionsvoraussetzung

$$r_{i+1} = \left[\left(1 + \frac{p_i}{q_i} - \frac{1}{q_i} A\right) R_i(A) - \frac{p_i}{q_i} R_{i-1}(A)\right] k \ .$$

Setzt man hierin für $A R_i(A)$ die aus (34) folgende Formel

$$A R_i(A) = - q_i R_{i+1}(A) + (p_i + q_i) R_i(A) - p_i R_{i-1}(A)$$

ein, so folgt in der Tat (37). Speziell ergibt sich

$$r_n = R_n(A)k \ .$$

Die Konfrontation mit (21) zeigt, daß beim alten Prozeß (8) und beim neuen Prozeß (36) die End-Residuen übereinstimmen, und somit fallen auch die Endpunkte zusammen. Damit ist unsere Behauptung bewiesen. Die Vorschrift (36) unterscheidet sich übrigens von einem Gradientenprozeß nur in dem zusätzlichen Term in Δx_{i-1}, der also die Vorgeschichte des Verfahrens berücksichtigt.

Unser neu definiertes Relaxationsverfahren vermeidet nun die genannten Nachteile. Erstens ist der Polygonzug durchaus eindeutig definiert. Von ausschlaggebender Bedeutung ist ferner die Tatsache, daß nun wegen (37) *nicht nur der Endpunkt ein Punkt bester Strategie ist, sondern auch jeder Zwischenpunkt x_i des Polygonzugs.* Ein Iterationsverfahren mit dieser Eigenschaft wollen wir ein *anfangskonvergentes* Verfahren nennen.

Satz 2. *Zu jeder Dichtefunktion $\varrho(\lambda)$ gehört ein anfangskonvergentes Relaxationsverfahren. Es wird gefunden, indem man die zur Dichtefunktion*

$\varrho(\lambda)$ *gehörige Schar orthogonaler Polynome konstruiert und mit Hilfe der zwischen ihnen bestehenden Rekursionsformeln* (34) *nach der Vorschrift* (36) *rechnet.*

Im Falle des Beispiels von Nr. III lautet diese Relaxationsvorschrift

$$\Delta x_i = \frac{1}{i+2}\left[4(i+1)r_i + i\,\Delta x_{i-1}\right] . \tag{39}$$

Zweitens funktioniert (36) auch für die diskontinuierlichen Fehlermasse. Um dies zu zeigen, berechnen wir die Koeffizienten in der Entwicklung (34):

$$q_i = -\frac{1}{N_{i+1}}\int_0^1 \lambda R_i(\lambda) R_{i+1}(\lambda)\varrho(\lambda)d\lambda\,, \quad p_i = -\frac{1}{N_{i-1}}\int_0^1 \lambda R_i(\lambda) R_{i-1}(\lambda)\varrho(\lambda)d\lambda,$$

$$p_i + q_i = \frac{1}{N_i}\int_0^1 \lambda R_i(\lambda)^2\varrho(\lambda)d\lambda\ .$$

Dabei ist N_i eine Abkürzung für die Norm

$$N_i = \int_0^1 R_i(\lambda)^2\varrho(\lambda)d\lambda\ . \tag{40}$$

Diesen Formeln geben wir die etwas einfachere Gestalt

$$q_i = \frac{1}{N_i}\int_0^1 \lambda R_i(\lambda)^2\varrho(\lambda)d\lambda - p_i\ , \qquad p_i = \frac{N_i}{N_{i-1}}\,q_{i-1}\ . \tag{41}$$

Für das Fehlermaß φ' von Formel (11) ergibt die zugehörige Dichte (25) unter Beachtung von (37) und (23)

$$\int_0^1 \lambda R_i(\lambda)^2\varrho(\lambda)d\lambda = \sum_{(j)} k_j^2 \int_0^1 \lambda R_i(\lambda)^2\,\delta(\lambda-\lambda_j)d\lambda$$
$$= \sum_{(j)} k_j^2\,\lambda_j R_i(\lambda_j)^2 = \sum_{(j)} \lambda_j r_{ij}^2 = (r_i, A r_i)$$

und analog

$$N_i = (r_i, r_i)\ ,$$

also

$$q_i = \frac{1}{N_i}(r_i, A r_i) - p_i\ , \qquad p_i = \frac{N_i}{N_{i-1}}\,q_{i-1}\ . \tag{42}$$

Der erste Teil von q_i ist der Rayleighsche Quotient des Residuums r_i. Der Algorithmus (36) (42) stimmt überein mit dem Verfahren, das in früheren Arbeiten als *n-Schritt-Verfahren* (*Stiefel* [5]), oder Methode der *konjugierten Gradienten* (*Hestenes, Stiefel* [8] speziell S. 423 Formeln (9 : 6)) bezeichnet wurde. Da dieses Verfahren an den zitierten Stellen ausführlich geschildert wurde und außerdem seither in verschiedenen

Richtungen verallgemeinert wurde (vgl. Hestenes [9], Householder [10]), formulieren wir nur das aus unseren Strategie-Überlegungen folgende Resultat.

Satz 3. *Die Methode der konjugierten Gradienten gibt einen kleineren Fehlervektor (im Sinne des A-Maßes von Nr. I) als jedes Gradientenverfahren, das aus höchstens gleichviel Schritten besteht.*

Benutzt man aber das Fehlermaß φ'' von (13), also die Quadratsumme der Residuen, so erhält man einen Algorithmus, der durch das folgende Formelsystem vollständig beschrieben wird.

$$r_0 = k \ , \qquad r_{i+1} = r_i + \Delta r_i \ , \qquad \Delta r_i = \frac{1}{q_i}(p_i \Delta r_{i-1} - A r_i) \ . \tag{43}$$

(Diese Formeln folgen aus (38) und (36).) Dabei ist

$$p_0 = 0 \ , \qquad q_i = \frac{1}{N_i}(A r_i, A r_i) - p_i \ , \qquad p_i = \frac{N_i}{N_{i-1}} q_{i-1} \ , \tag{44}$$

wobei

$$N_i = (r_i, A r_i) \ .$$

Der erste Teil von q_i ist Schwarzscher Quotient des Residuenvektors. Parallel damit oder nachträglich berechnet man die Näherungspunkte

$$x_0 = 0 \ , \qquad x_{i+1} = x_i + \Delta x_i \ , \qquad \Delta x_i = \frac{1}{q_i}(r_i + p_i \Delta x_{i-1}) \ . \tag{45}$$

Satz 4. *Der soeben beschriebene Algorithmus liefert eine kleinere Quadratsumme der Residuen als jedes Gradientenverfahren, das aus höchstens gleichviel Schritten besteht.*

Die Verwendung des Fehlermaßes φ von Formel (10) ist jedoch auch hier nicht möglich, da in den Formeln die inverse Matrix A^{-1} auftaucht.

Das Verfahren (43) bis (45) kann als Weiterentwicklung einer Methode von *Lanczos* [11] angesehen werden.

V. Wir haben in Satz 2 unsere Iterationen als „anfangskonvergent" bezeichnet. Der Ordnung halber soll jetzt noch die Konvergenz im streng mathematischen Sinn untersucht werden. Dazu brauchen wir eine Information über den Abbau des Fehlermaßes. Wir verwenden zunächst, daß

$$\frac{R_{i+1} - R_i}{\lambda}$$

wegen (32) ein Polynom vom Grad i ist. Da R_{i+1} auf allen Polynomen

niedrigeren Grades orthogonal steht, folgt

$$\int \frac{R_{i+1} - R_i}{\lambda} R_{i+1} \varrho d\lambda = 0 , \quad \int \frac{R_i R_{i+1}}{\lambda} \varrho d\lambda = \int \frac{R_{i+1}^2}{\lambda} \varrho d\lambda . \tag{46}$$

Sodann schreiben wir die Rekursionsformel (34) in der Gestalt

$$R_i = - q_i \frac{R_{i+1}}{\lambda} + q_i \frac{R_i}{\lambda} + p_i \frac{R_i - R_{i-1}}{\lambda} .$$

Multiplikation mit R_i und Integration ergibt

$$\int R_i^2 \varrho d\lambda = - q_i \int \frac{R_{i+1} R_i}{\lambda} \varrho d\lambda + q_i \int \frac{R_i^2}{\lambda} \varrho d\lambda$$

und wegen (46)

$$\int R_i^2 \varrho d\lambda = - q_i \int \frac{R_{i+1}^2}{\lambda} \varrho d\lambda + q_i \int \frac{R_i^2}{\lambda} \varrho d\lambda .$$

Aus der Definition (24) des Fehlermaßes und aus (40) folgt

$$\psi_{i+1} - \psi_i = - \frac{1}{q_i} N_i . \tag{47}$$

Diese Formel kann zum Beispiel dazu dienen, beim Algorithmus das Fehlermaß leicht mitzuberechnen. Wir benutzen sie aber jetzt zur Herleitung einiger Ungleichungen. Da ψ bei jedem Schritt sicher verringert wird, folgt zunächst $q_i > 0$. Aus der zweiten Gleichung (41) schließt man $p_i > 0$ und aus der ersten

$$q_i < \frac{1}{N_i} \int_0^1 \lambda R_i^2 \varrho d\lambda \leqslant \frac{1}{N_i} \int_0^1 R_i^2 \varrho d\lambda = 1 .$$

Endgültig

$$0 < q_i < 1 . \tag{48}$$

Die Zahlen ψ_i bilden eine monoton abnehmende Folge positiver Zahlen, es existiert also $\lim \psi_i$ und somit liefert (47)

$$\lim \frac{1}{q_i} N_i = - \lim (\psi_{i+1} - \psi_i) = 0 .$$

Wegen (48) folgt auch

$$\lim N_i = 0 .$$

Satz 5. *Das Relaxationsverfahren von Satz 2 konvergiert im Mittel, das heißt*

$$\int_0^1 R_i(\lambda)^2 \varrho(\lambda) d\lambda \to 0 .$$

Das Verfahren der konjugierten Gradienten und seine Variante (43) bis (45) nehmen hinsichtlich der Konvergenz eine Sonderstellung ein. Aus der Orthogonalität der Residuenpolynome

$$\int_0^1 R_i(\lambda) R_k(\lambda) \varrho(\lambda) d\lambda = 0 \quad \text{für} \quad i \neq k$$

folgt nämlich beziehlich beim Einsetzen der Dichten (25) (26)

$$(r_i, r_k) = 0 \ , \quad (r_i, A r_k) = 0 \quad \text{für} \quad i \neq k \ .$$

Die Residuenvektoren bilden also ein Orthogonalsystem im gewöhnlichen Sinn bzw. im Sinn der A-Metrik. Da es nur endlich viele von Null verschiedene Orthogonalvektoren geben kann, muß unser Relaxationsprozeß nach endlich vielen Schritten zu einem Null-Residuum, das heißt zum Lösungspunkt des Gleichungssystems, führen. Die Anzahl der Schritte ist höchstens gleich der Anzahl der gegebenen Gleichungen (1).

Für die Praxis ist dieses Resultat ohne große Bedeutung, da die Anzahl der Schritte, für welche diese beiden immerhin komplizierten Relaxationsverfahren durchzurechnen sind, meistens viel zu groß ist und sich oft numerische Instabilitäten zeigen. Wir benutzen vielmehr folgende Tatsache. Wenn der Vektor k in (1) in einem gegenüber A invarianten Teilraum R' des totalen Raumes R liegt, so führen beide Verfahren nach höchstens soviel Schritten zur exakten Lösung, wie die Dimension von R' angibt.

Sei nun ein *schlecht konditioniertes* Gleichungssystem gegeben, dann kann man nach einem Vorschlag von *H. Rutishauser*, der auf der Idee des „purified vector" von *Lanczos* [11] beruht, folgendermaßen vorgehen. Man wähle eine Zahl $\varepsilon > 0$, die zwar nicht untere Grenze der Eigenwerte zu sein braucht, aber doch nur wenige Eigenwerte links von sich läßt. Zunächst werden nun nach einem einfachen Relaxationsverfahren (etwa nach dem am Ende von Nr. III beschriebenen) die Residuen beseitigt, die zu den Eigenwerten im Intervall von ε bis 1 gehören. Der verbleibende Residuenvektor liegt in demjenigen A-invarianten Teilraum, der aufgespannt wird von den Eigenvektoren, die zu den Eigenwerten links von ε gehören. Fährt man nun nach der feineren Methode der konjugierten Gradienten oder nach dem Algorithmus (43) bis (45) fort, so wird der exakte Lösungspunkt nach höchstens sovielen Schritten erreicht, wie die Anzahl der links von ε gelassenen Eigenwerte angibt. Ist diese nicht zu groß, so werden auch die genannten Algorithmen stabil bleiben.

Mitunter ist es sogar angezeigt, die links von ε liegenden Eigenwerte zu berechnen. Dafür wird mit Vorteil der Quotienten-Differenzen-Algo-

rithmus von *Rutishauser* [12] verwendet, für den sowohl die Relaxation nach konjugierten Gradienten als auch die Vorschrift (43) bis (45) die nötigen Grundlagen liefern. Man relaxiert dann weiter nach dem Gradientenverfahren, (8), indem man für die q_i diese Eigenwerte nimmt.

Zum Abschluß sei noch bemerkt, daß unsere ganze Theorie bis und mit Satz 5 auch dann gilt, wenn A ein symmetrischer und *vollstetiger Operator* im Hilbert-Raum ist, dessen Spektrum im Intervall von 0 bis 1 der λ-Achse liegt, ohne sich im Nullpunkt zu häufen. Auch die eben geschilderte kombinierte Relaxationsmethode bleibt durchführbar, falls links von der gewählten Zahl ε nur endlich viele Eigenwerte liegen.

VI. Wir kehren zurück zu stetigen Dichtefunktionen $\varrho(\lambda)$. Hat man keine andere Information über die Eigenwerte von A als daß sie im Intervall 0 bis 1 liegen, so muß man wählen

$$\varrho(\lambda) > 0 \quad \text{für} \quad 0 < \lambda < 1 \ ,$$

damit alle Residuen angegriffen werden. Andererseits muß $\varrho(0) = 0$ sein, damit das Integral

$$\int_0^1 \frac{\varrho(\lambda)}{\lambda} d\lambda$$

existieren kann, wie in Nr. III vorausgesetzt wurde. Ferner kann $\varrho(1) = 0$ gewählt werden, weil oberhalb 1 keine Residuen liquidiert werden müssen. Aus diesen Gründen werden wir $\varrho(\lambda)$ in der Form schreiben

$$\varrho(\lambda) = \lambda^\alpha (1 - \lambda)^\beta f(\lambda) \ , \quad \alpha > 0 \ , \quad \beta \geqslant -\tfrac{1}{2} \ . \tag{49}$$

Dabei sei etwa $f(\lambda)$ eine im Intervall $0 \leqslant \lambda \leqslant 1$ zweimal stetig differenzierbare Funktion und

$$f(\lambda) > 0 \quad \text{für} \quad 0 \leqslant \lambda \leqslant 1 \ . \tag{50}$$

Satz 6. *Unter diesen Voraussetzungen für die Dichtefunktion konvergiert die Relaxation von Satz 2, das heißt, die nach* (36) *konstruierte Folge von Näherungspunkten x_i strebt gegen den Lösungspunkt des gegebenen Gleichungssystems.*

Beweis. Nach Szegö [3], Seite 291, Formel (12.1.12) gilt für die Orthogonalpolynome $R_i(\lambda)$ zur gegebenen Dichte die asymptotische Formel für große Werte von i

$$R_i(\lambda) \sim \sqrt{\frac{2N_i}{\pi}} \frac{\lambda^{-\frac{\alpha}{2}-\frac{1}{4}} (1-\lambda)^{-\frac{\beta}{2}-\frac{1}{4}}}{\sqrt{f(\lambda)}} \cos(i\varphi + \gamma) \ , \quad 0 < \lambda < 1 \ , \tag{51}$$

wobei N_i wieder die Norm und γ eine hier nicht weiter interessierende Phase bedeutet und wieder $\cos\varphi = 1 - 2\lambda$ gesetzt wurde. Wegen Satz 5 strebt die Norm gegen Null, und daher folgt

$$R_i(\lambda) \to 0 \quad \text{für} \quad 0 < \lambda < 1 \ . \tag{52}$$

Aus (23)

$$r_{ij} = R_i(\lambda_j) k_j \to 0 \quad \text{für} \quad i \to \infty \text{ und alle } j \ .$$

Damit ist der Satz bewiesen.

Wegen $R(0) = 1$ ist natürlich die Konvergenz (52) ungleichmäßig in bezug auf λ; Eigenwerte, die nahe bei $\lambda = 0$ liegen, machen also auch hier Schwierigkeiten. Die Aussage (51) bedeutet noch, daß $R_i(\lambda)$ eine Schwingung ist mit der *Amplitudenfunktion*

$$a_i(\lambda) \sim \sqrt{\frac{2N_i}{\pi}} \frac{\lambda^{-\frac{\alpha}{2}-\frac{1}{4}} (1-\lambda)^{-\frac{\beta}{2}-\frac{1}{4}}}{\sqrt{f(\lambda)}} \ . \tag{53}$$

Durch Wahl von α, β; $f(\lambda)$ kann man diese Amplitude beliebig formen, um so der anfänglichen Residuenverteilung Rechnung zu tragen.

Wir diskutieren weiter die einfachste mögliche Wahl

$$f(\lambda) = 1 \quad \text{also} \quad \varrho(\lambda) = \lambda^\alpha (1-\lambda)^\beta \ . \tag{54}$$

Die Polynome $R_n(\lambda)$ werden dann *hypergeometrische* (Jacobische) Polynome, die wir bezeichnen mit

$$R_n(\alpha, \beta; \lambda) = F(-n, n+\alpha+\beta+1, \alpha+1; \lambda) \ , \tag{55}$$

wobei F die Gaußsche hypergeometrische Funktion bedeutet. Die Nebenbedingung (22) ist erfüllt. Die zwischen ihnen bestehende Rekursionsformel (Szegö [3], Seite 71) liefert nach (36) folgende einfache Relaxationsvorschrift, die wir *hypergeometrische Relaxation* nennen.

$$x_0 = 0 \ , \quad x_{i+1} = x_i + \Delta x_i \ , \quad r_i = k - A x_i \ ,$$

$$\Delta x_i = \frac{1}{q_i}(r_i + p_i \Delta x_{i-1}) \ ,$$

wobei

$$q_i = \tfrac{1}{4}\left\{1 + \frac{(\alpha-\beta+1)(\alpha+\beta+1)}{2i+\alpha+\beta+1} - \frac{(\alpha-\beta)(\alpha+\beta)}{2i+\alpha+\beta+2}\right\} , \tag{56}$$

$$p_i = \tfrac{1}{4}\left\{1 - \frac{(\alpha-\beta+1)(\alpha+\beta+1)}{2i+\alpha+\beta+1} + \frac{(\alpha-\beta)(\alpha+\beta)}{2i+\alpha+\beta}\right\} .$$

Der Relaxationsprozeß hängt noch von den beiden Parametern α, β ab, deren Bedeutung jetzt abgeklärt werden soll. Setzt man in (53) eine geeignete asymptotische Annäherung für die Norm der hypergeometrischen Polynome ein, so erhält man zunächst die Amplitudenfunktion

$$a_n(\lambda) \sim \frac{\alpha!}{\sqrt{\pi}\, n^{\alpha+\frac{1}{2}}} \lambda^{-\frac{\alpha}{2}-\frac{1}{4}} (1-\lambda)^{-\frac{\beta}{2}-\frac{1}{4}}, \tag{57}$$

die nur für $0<\lambda<1$ gültig ist. Diese Amplitude nimmt ab bis zum Punkt

$$\lambda = \frac{2\alpha+1}{2(\alpha+\beta+1)} \tag{58}$$

und dann wieder zu. Für $\beta = -\frac{1}{2}$ befindet sich dieser Punkt am rechten Ende des Intervalls von 0 bis 1, für $\beta<\alpha$ rechts von der Mitte und für $\beta>\alpha$ links von der Mitte.

Bedeutung von α. Um diese abzuklären, brauchen wir feinere asymptotische Formeln. Da wir insbesondere die Anwendung der hypergeometrischen Relaxation auf schlecht konditionierte Systeme im Auge haben, begnügen wir uns damit, das Verhalten von $R_n(\alpha, \beta; \lambda)$ für kleine λ und große n anzugeben. Setzen wir

$$x = 2n\sqrt{\lambda}\,, \tag{59}$$

so gilt für festes x und $n \to \infty$ nach Szegö [3], Seite 186, Formel (8.1.1)

$$R_n(\alpha, \beta; \lambda) \sim \Lambda_\alpha(x)\,. \tag{60}$$

Dabei ist $\Lambda_\alpha(x)$ die in *Jahnke-Emde* [13] tabellierte und dort in Fig. 96 graphisch dargestellte Funktion

$$\Lambda_\alpha(x) = \frac{2^\alpha \alpha!}{x^\alpha} J_\alpha(x)\,,$$

wobei $J_\alpha(x)$ Besselsche Funktion erster Art ist.

Die Formel (60) zeigt zunächst, daß der Parameter β für kleine λ ohne Einfluß auf die Konvergenzverhältnisse ist. Die Funktion $\Lambda_\alpha(x)$ beginnt mit $\Lambda_\alpha(0) = 1$, was der Bedingung (22) entspricht. Für wachsendes x hat sie die Gestalt einer gedämpften Schwingung; in der folgenden Tabelle ist die Lage und Höhe des ersten Minimums für einige Werte von α angegeben.

α	x	$\Lambda_\alpha(x)$
0,5	4,49	— 0,217
1	5,14	— 0,132
1,5	5,76	— 0,086
2	6,38	— 0,059
3	7,59	— 0,029
4	8,77	— 0,016
5	9,94	— 0,009

(61)

Relaxiert man also zum Beispiel mit $\alpha = 5$, so liegt das Minimum von $R_n(\lambda)$ nach (59) bei

$$\lambda = \left(\frac{4{,}97}{n}\right)^2$$

und hat für alle n den Wert $-0{,}009$. (Diese Aussagen gelten nur für große n.) Es wandert also mit wachsendem n in der Richtung der kleineren Eigenwerte unter Beibehaltung seiner Höhe. Man wird nur so lange relaxieren, bis dieses Wellental über den kleinsten Eigenwert hinwegläuft, was man daran erkennt, daß alle Residuen auf 0,9 Prozent ihrer anfänglichen Werte gesunken sind. Sollen die Residuen nur auf 10 Prozent abgebaut werden, so verwendet man etwa $\alpha = 1{,}5$. Die kleinen Eigenwerte werden dann früher erreicht.

Bedeutung von β. Die Wahl von β äußert sich in erster Linie im Verhalten der Relaxation im rechten Teil des Intervalls $0 < \lambda < 1$. Wählt man $\beta = -\frac{1}{2}$, so nimmt nach (57) die Amplitude monoton mit λ ab, die zu höheren Eigenwerten gehörigen Residuen werden also am stärksten angegriffen. Löst man ein Randwertproblem der mathematischen Physik mit Differenzenrechnung in einem Punktgitter, so bedeutet dies, daß die Residuenverteilung *geglättet* wird. Dies bleibt richtig solange $\beta < \alpha$ ist, wobei die Residuen an der Stelle (58) am stärksten angegriffen werden. Für $\beta > \alpha$ jedoch gibt es eine Stelle im Intervall, für welche die Amplitude (57) den Wert 1 hat. Rechts davon ist sie größer als 1; in dieser Gegend werden also die Residuen aufgeschaukelt, das heißt die Residuenverteilung wird *aufgerauht*. Diese Rauhigkeit verschwindet wegen Satz 6 wieder mit wachsendem n, indem die kritische Stelle immer weiter nach rechts rückt. Eine ähnliche Aufrauhung tritt ein, wenn man bei elementareren Relaxationsmethoden die sogenannte Überrelaxation anwendet. Wir nennen daher β den *Parameter* der *Überrelaxation*.

Um die Wirkung einer Überrelaxation in der Gegend von $\lambda = 0$ abzuklären, genügen asymptotische Betrachtungen nicht mehr. In der folgenden Tabelle ist die Lage des ersten Minimums von $R_n(\alpha, \beta; \lambda)$ in den beiden typischen Fällen $\alpha = 0{,}5$, $\beta = 0{,}5$ und $\alpha = 0{,}5$, $\beta = 10{,}5$ angegeben.

	$\beta = 0{,}5$		$\beta = 10{,}5$	
n	λ	$R_n(\lambda)$	λ	$R_n(\lambda)$
2	0,500	— 0,333	0,167	— 0,555
4	0,194	— 0,250	0,075	— 0,330
6	0,101	— 0,233	0,046	— 0,278
8	0,062	— 0,227	0,029	— 0,258
10	0,041	— 0,223	0,022	— 0,243
12	0,030	— 0,222	0,018	— 0,230
12 (as)	0,035	— 0,217	0,035	— 0,217

(62)

Die letzte Zeile enthält die nach Tabelle (61) berechnete asymptotische Lage des Minimums für $n = 12$. Man erkennt, daß für mäßige Werte von n die Überrelaxation besser in das Gebiet der kleinen Eigenwerte eindringt. Die mit ihr verbundene Aufrauhung der Residuen kann bei Randwertaufgaben oft in Kauf genommen werden. In vielen Fällen ist dann nämlich die anfängliche Residuenverteilung glatt, so daß am Anfang die zu den höheren Eigenwerten gehörigen Residuenkomponenten klein sind und ihre Aufschauklung nichts schadet. Falls man also die Überrelaxation anwenden kann, bewirkt sie eine Beschleunigung der Anfangskonvergenz.

VII. Die verschiedenen hier angeregten Methoden wurden experimentell nachgeprüft am Beispiel der Torsion eines quadratischen Querschnitts G durch Lösung der partiellen Differentialgleichung

$$\Delta f = -1 \,, \qquad f = 0 \quad \text{am Rande} \,. \tag{63}$$

Dabei ist Δ der Laplacesche Differentialoperator. Es wurde ein Gitter von 10×10 Teilquadraten in G gelegt, so daß f in 81 inneren Gitterpunkten bestimmt werden mußte. Als Annäherung des Δ-Operators in (63) wurde die Summe der zweiten partiellen Differenzen in den beiden Koordinatenrichtungen genommen. Die Relaxation geschah ohne Benutzung der Symmetrien des Quadrats, so daß die folgenden Ergeb-

nisse auch informativ sind für ein allgemeiner gestaltetes Gebiet mit etwa 80 Gitterpunkten. Ausgangsfunktion für die Relaxation ist in allen Fällen die in G identisch verschwindende Funktion.

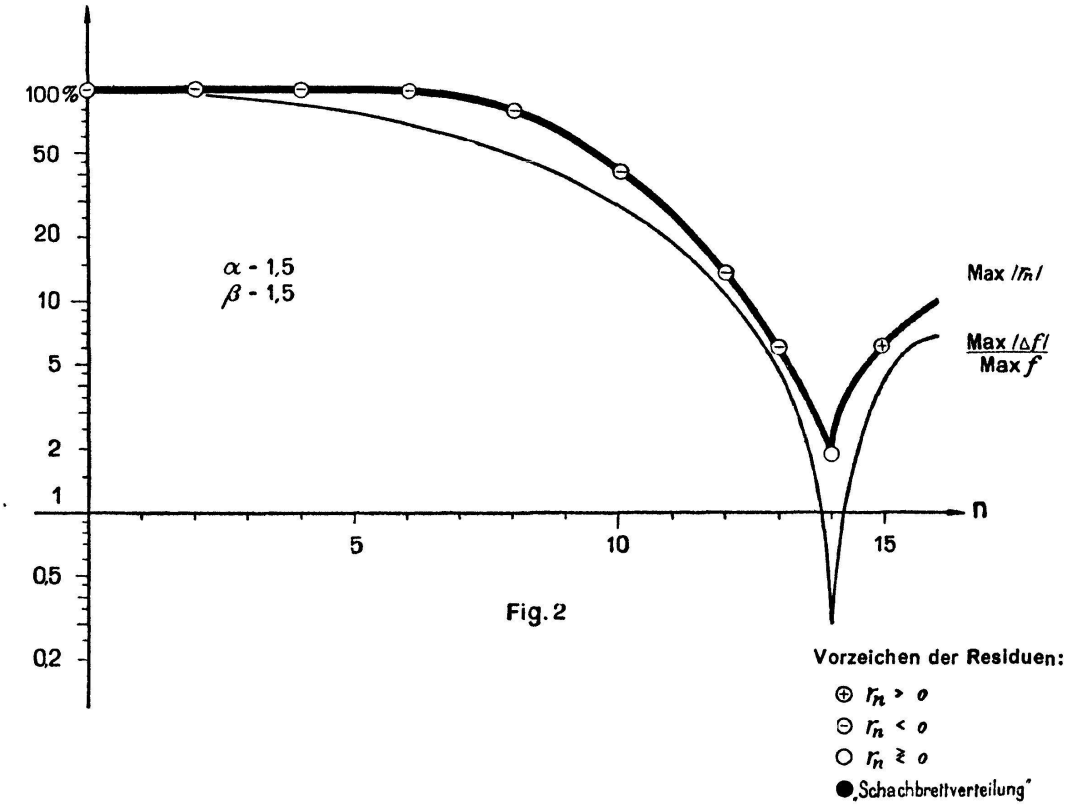

Fig. 2

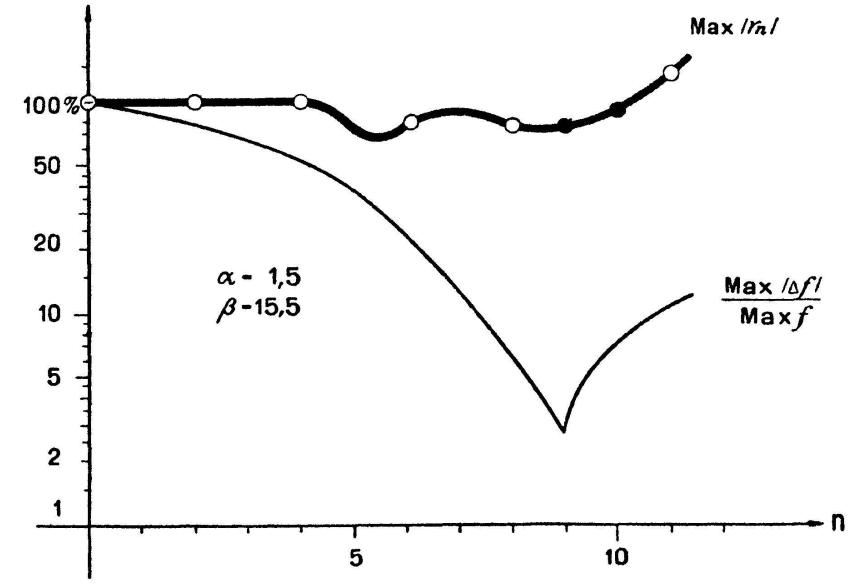

Fig. 3

Die Figuren 2 und 3 zeigen hypergeometrische Relaxationen. Es sind jeweilen als Funktion der Schrittzahl n aufgetragen:

1. Das maximale Residuum r_n in G in Prozenten des Ausgangsresiduums, welches in allen Gitterpunkten -1 ist.

2. Der maximale Fehler der betreffenden Näherungslösung gegenüber der exakten Lösung in Prozenten des maximalen exakten Wertes von f, der natürlich in der Mitte des Quadrats angenommen wird. Zur Beurteilung der Rauhigkeit der Residuen ist außerdem angegeben, ob die Residuen überall in G dasselbe Vorzeichen haben oder nicht, und es wurde speziell hervorgehoben, wenn die sogenannte Schachbrettverteilung eintritt, das heißt, wenn jede Gitterstrecke an ihren Endpunkten Residuen verschiedenen Vorzeichens trägt.

Figur 2 zeigt eine normale Relaxation mit $\alpha = 1{,}5$, $\beta = 1{,}5$. Die Residuen werden abgebaut bis zum 14. Schritt unter Beibehaltung des negativen Vorzeichens in allen Residuen. Beim 14. Schritt schlagen die Vorzeichen um. Dies ist also der in Nr. VI erwähnte Moment, wo das Wellental von $R_n(\lambda)$ über den kleinsten Eigenwert streicht. Bei diesem Schritt sollte man aufhören.

Figur 3 veranschaulicht demgegenüber eine typische Überrelaxation mit $\alpha = 1{,}5$, $\beta = 15{,}5$. Die Residuen erhalten schon nach dem ersten Schritt verschiedene Vorzeichen in G und beim 9. Schritt tritt die Schachbrettverteilung ein. Das maximale Residuum nimmt überhaupt nicht ab. Trotzdem erhält man mit 9 Schritten ein genaueres Resultat als in Figur 2, indem der maximale Fehler im Funktionswert auf etwa 3 Prozent abgesunken ist. Man kann also den Eintritt der Schachbrettverteilung als Signal dafür werten, daß die Überrelaxation abzubrechen ist.

Dasselbe Beispiel wurde auch nach der am Ende von Nr. V geschilderten *kombinierten Methode* durchgerechnet. Nach Wahl der unteren Grenze $\varepsilon = 0{,}25$ wurden zunächst im Intervall

$$0{,}25 \leqslant \lambda \leqslant 1 \tag{64}$$

die Residuen genügend liquidiert, indem 11 Schritte des rohen Gradientenprozesses (8) ausgeführt wurden. Dabei wählte man für die Faktoren q_i Zahlen im Intervall (64) mit einer gewissen Verdichtung an den Enden des Intervalls. Anschließend folgten 2 Schritte des Prozesses (43) (44). Die folgende Tabelle gibt den Stand von Residuen und Fehlern vor und während dieses zweiten Teils der Rechnung

n	Max $\lvert r_n \rvert$	$\dfrac{\text{Max} \lvert \Delta f \rvert}{\text{Max } f}$
0	88%	70%
1	23%	5,1%
2	0,2%	0,014%

was genügend für die Wirksamkeit des kombinierten Verfahrens spricht. An Stelle des Gradientenverfahrens (8) könnte man die anfängliche Glättung auch mit hypergeometrischer Relaxation ausführen.

LITERATURVERZEICHNIS

[1] *Magnus R. Hestenes:* The solution of linear equations by minimization, U. S. Department of Commerce, National Bureau of Standards, NAML Report 52–45, 1951.

[2] *Marvin L. Stein:* Gradient methods in the solution of systems of linear equations, J. of research of the National Bureau of Standards, vol. 48, 1952, p. 407–413.

[3] *G. Szegö:* Orthogonal polynomials, Amer. Math. Soc. colloquium publications, vol. XXIII, 1939.

[4] *C. Lanczos:* Tables of Chebyshev Polynomials, National Bureau of Standards, applied Math. series 9, 1952.

[5] *E. Stiefel:* Über einige Methoden der Relaxationsrechnung, Z. angew. Math. Physik, vol. III, 1952, S. 2–33.

[6] *C. Lanczos:* Chebyshev polynomials in the solution of largescale linear systems, Proc. of the assoc. for computing machinery, Toronto meeting 1952.

[7] *D. Young:* On Richardson's method for solving linear systems with positive definite matrices, J. Math. Phys. (1954) 244–255.

[8] *M. Hestenes and E. Stiefel:* Methods of conjugate gradients for solving linear systems, J. of research of the National Bureau of Standards, vol. 49, 1952, p. 409–435.

[9] *M. Hestenes:* The conjugate gradient method for solving linear systems, Report of the Institute for Numerical Analysis, Los Angeles, Jan. 11, 1954.

[10] *A. S. Householder:* On solving linear algebraic systems, Report of the Oak Ridge National Laboratory, Sep. 11, 1954.

[11] *C. Lanczos:* Solution of systems of linear equations by minimized iterations, J. of research of the National Bureau of Standards, vol. 49, 1952, p. 33–53.

[12] *H. Rutishauser:* Bestimmung der Eigenwerte und Eigenvektoren einer Matrix mit Hilfe des Quotienten-Differenzen-Algorithmus. (Erscheint demnächst in Z. angew. Math. Physik.)

[13] *Jahnke-Emde:* Tafeln höherer Funktionen, 4. Aufl., Leipzig, B. G. Teubner, S. 180–189.

(Eingegangen den 9. Dezember 1954.)