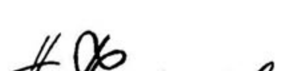
Beiträge zur Numerischen Mathematik 4



Beiträge
zur Numerischen Mathematik 4

Herrn Prof. Dr.-Ing. habil. Dr. techn. h. c. Helmut Heinrich
zum 70. Geburtstag gewidmet

Herausgegeben von
Frieder Kuhnert und Jochen W. Schmidt

R. Oldenbourg Verlag München Wien 1975

Lizenzausgabe für den R. Oldenbourg Verlag München Wien
Printed in the German Democratic Republic
Lizenz-Nr. 206 · 435/111/75
Gesamtherstellung: VEB Druckhaus „Maxim Gorki", 74 Altenburg
LSV 1035

ISBN: 3-486-20131-x

Helmut Heinrich 70 Jahre alt

Helmut Heinrich wurde am 5. September 1904 in Gottesberg (heute Boguszów) geboren. Im Jahre 1923 legte er am Gymnasium in Schweidnitz (heute Świdnica) die Reifeprüfung ab und studierte von 1924 bis 1928 an der Technischen Hochschule Breslau (heute Wrocław) Mathematik. Seine Absicht, nach Ablegung der Diplomhauptprüfung als Mathematiker in der Praxis, und zwar in der Flugzeugindustrie zu arbeiten, konnte er wegen der schlechten wirtschaftlichen Verhältnisse nicht verwirklichen. Er erhielt die Möglichkeit, zunächst vertretungsweise in Glogau (heute Głogów) Mathematik und Physik zu unterrichten, und wurde nach Ablegung der ersten (1929) und zweiten (1931) Staatsprüfung für das Lehramt an höheren Schulen als Studienreferendar bzw. Studienassessor in den Schuldienst aufgenommen. Aber auch diese Tätigkeit sollte keine dauernde Lösung bleiben. Eine Notverordnung zwang Heinrich, seine Stellung aufzugeben. Er fand jedoch einen Ausweg in dieser schwierigen Lage, indem er unentgeltlich als Lehrer weiterarbeitete und seinen Lebensunterhalt durch eine pädagogische Nebentätigkeit verdiente.

Seine ursprünglich gehegte Absicht, wissenschaftlich zu arbeiten, mag durch diese Umstände wieder verstärkt hervorgetreten sein. Er nahm die Verbindung mit seiner Ausbildungsstätte, dem von Prof. Dr. Werner Schmeidler geleiteten Lehrstuhl für höhere Mathematik der TH Breslau wieder auf, arbeitete dort in den Ferien als Hilfsassistent und promovierte im Jahre 1933 zum Dr.-Ing. mit der Arbeit „Über die Bedeutung der Pfeilstellung eines Tragflügels“, einem zu dieser Zeit sehr aktuellen Thema. Von 1933 bis 1936 betreute er als Professor für Mathematik an der Staatlichen Chinesischen Tung-Chi-Universität Woosung die mathematische Ausbildung der Ingenieur-Studenten. Nach seiner Rückkehr erhielt er Lehraufträge an der TH Breslau. Im Jahre 1937 habilitierte er sich mit der Arbeit „Ein abbildungsgeometrisches Verfahren zur Darstellung von Richtungsfeldern ...“ und wurde im Jahre 1938 zum Dozenten für reine und angewandte Mathematik ernannt.

In den folgenden Jahren entfaltete Heinrich eine umfangreiche Lehrtätigkeit, die sich auf beinahe alle Teilgebiete der reellen und komplexen Analysis sowie auf die praktische Mathematik erstreckte. Gleichermaßen bedeutsam war aber für ihn, daß er nun günstigere Möglichkeiten für eine wissenschaftliche Tätigkeit vorfand, als sie sich ihm bisher geboten hatten. Eine Durchmusterung der von 1932 bis 1941 entstandenen Arbeiten verrät den wissenschaftlichen Arbeitsstil Heinrichs. Auf den ersten Blick fällt zunächst eine große Vielseitigkeit auf. Für Heinrich gibt es nicht

„das große Problem seines Lebens“, neben dem nichts anderes mathematisch Interessantes existiert. Ihn reizt jedes Problem, das er mit seinen mathematischen Fähigkeiten der Lösung ein Stück näher bringen kann, und es ist überall das Bestreben zu spüren, ein in der Praxis auch wirklich anwendbares Ergebnis zu finden und die Lösung so zu formulieren, daß der Sachverhalt einem möglichst breiten Leserkreis verständlich ist.

Auf verschiedenen Gebieten hat Heinrich damit Pionierarbeit geleistet. Zum Beispiel dürfte die erste, im Jahre 1932 erschienene kleine Arbeit über die Frage „Ist Schrägüberschreiten der Fahrbahn gefährlich?“ zu den ersten quantitativen Untersuchungen zur Verkehrssicherheit zählen. In weiteren Arbeiten wird mit großem Erfolg eine Verbindung zwischen mathematischen Gebieten ausgenutzt, deren Zusammenhang zunächst nicht ohne weiteres erkennbar war, nämlich der graphischen Lösung von gewöhnlichen Differentialgleichungen und der Nomographie und ihrer theoretischen Grundlage, der Abbildungsgeometrie. In der heutigen Zeit, in der immer mehr audiovisuelle Lehrmittel mit großem Erfolg eingesetzt werden, erscheint bemerkenswert, daß Heinrich bereits in den Jahren 1937/38 drei Filme für den mathematischen Hochschulunterricht schuf, in denen mit Hilfe eines einfachen Trickverfahrens die durch konforme Abbildung vermittelten Zusammenhänge für einige spezielle Funktionen veranschaulicht wurden. Bei der Beschäftigung mit numerischen Verfahren erkannte Heinrich klar, daß solche Verfahren künftig nicht nur vom Mathematiker benutzt werden und daß eindeutige Rechenvorschriften bereitgestellt werden müssen. Mit den damals zur Verfügung stehenden Rechenhilfsmitteln war dieses Ziel am besten durch die Angabe eines Rechenschemas zu erreichen; derartige Schemata wurden von ihm zum Rechnen mit Polynomen und zur Lösung von Polynomgleichungen vierten Grades angegeben. Hier handelt es sich offenbar um eine Vorstufe der heute so aktuellen algorithmischen Betrachtungsweise.

Im Zusammenhang mit der militärischen und politischen Entwicklung zum Ende des zweiten Weltkrieges ergaben sich auch für Heinrich einschneidende Veränderungen. An der Technischen Hochschule Dresden fand er — zunächst vorübergehend — einen neuen Arbeitsplatz. Dann nahm er eine Stellung in der Industrie an und war anschließend von Oktober 1946 bis Juni 1954 als Spezialist in der Sowjetunion tätig. Über die zwischen 1941 und 1954 erzielten Ergebnisse liegen aus verständlichen Gründen nur wenig Veröffentlichungen vor.

Nach seiner Rückkehr entschied sich Heinrich für ein Wirken in der Deutschen Demokratischen Republik. Mit der Ernennung zum Professor für Sondergebiete der angewandten Mathematik an der Technischen Hochschule Dresden eröffneten sich ihm ausgezeichnete Arbeitsmöglichkeiten. Er fand in Friedrich Adolf Willers einen verständnisvollen älteren Kollegen und darüber hinaus eine Anzahl jüngerer Mitarbeiter, mit denen ihn der gleiche Forschungsgegenstand und die gleiche Grundeinstellung zu Fragen der angewandten Mathematik verbanden. Bereits damals rechtfertigte er in seinen Vorlesungen den Ruf großen didaktischen Geschicks, der ihm vorausging. Er verstand es, auch schwierige Stoffgebiete so flüssig und klar gegliedert darzustellen, daß die Hörer gefesselt wurden und daß auch komplizierte Zusammenhänge verständlicher wurden.

Sehr bald arbeitete Heinrich tatkräftig an der Neugestaltung des Hochschulwesens in der DDR mit. Die Ernennung zum Mitglied des wissenschaftlichen Beirats für Leichtbau beim Staatssekretariat für Hochschulwesen im Jahre 1955 war der Auftakt für eine ganze Reihe von ehrenvollen Berufungen in beratende Gremien des Hochschulwesens, der Volksbildung und des Verlagswesens. Das Ausmaß des fördernden Einflusses, den HEINRICH in diesem Zusammenhang ausübte, kann nicht hoch genug eingeschätzt werden. Das liegt daran, daß er, wie auch sonst, sich voll für seine Arbeit engagierte und es nicht mit der Teilnahme an Sitzungen bewenden ließ; die dort behandelten Probleme bewegten ihn auch weiterhin und veranlaßten ihn zu den verschiedensten Aktivitäten, wie Umfragen, Materialsammlungen u. a. Er ließ sich auch nicht entmutigen, wenn er nicht sofort Gehör fand, sondern vertrat weiterhin den von ihm als richtig erkannten Standpunkt. Durch diese klare Haltung erwarb er sich die Achtung und das Vertrauen von verantwortlichen Mitarbeitern des Staatsapparates. Die Auszeichnung mit dem Vaterländischen Verdienstorden im Jahre 1961 war ein Dank des Staates auch für diese Seite seiner Tätigkeit.

Bereits im Jahre 1956 wurde HEINRICH als Nachfolger von WILLERS mit der Leitung des Instituts für Angewandte Mathematik betraut. Vor ihm stand die Aufgabe, das Institut weiter auszubauen und es in die Lage zu versetzen, den steigenden Anforderungen in Forschung und Lehre gerecht zu werden. In der Forschung sowie in der Ausbildung der Mathematiker galt es vor allem, die bisher eingeschlagene Richtung weiterzuverfolgen und dabei die rasche Entwicklung der elektronischen Rechentechnik zu berücksichtigen. Um die hohe Verantwortung, die mit der Lösung dieser Aufgabe verbunden war, richtig einschätzen zu können, muß man wissen, daß in den ersten Jahren die TH Dresden die Entwicklung der numerischen Mathematik in der DDR weitgehend allein zu tragen hatte.

Das Niveau der Mathematik-Ausbildung der Naturwissenschaftler, Ingenieure und Ökonomen galt es gleichfalls zu heben. Hierfür hat sich HEINRICH mit besonderer Energie eingesetzt, war es doch schon von jeher sein Anliegen, die Anwendbarkeit der Mathematik zu demonstrieren; in diesem Zusammenhang dürfte interessieren, daß er vor der Aufnahme seines Mathematik-Studiums in einer Maschinenfabrik arbeitete und auch ein Semester Maschinenbau studierte. Er setzte sich dafür ein, daß den Studenten die Mathematik in Form und Umfang derart dargeboten wird, daß sie für ihr weiteres Studium und die Arbeit in der Praxis von größtmöglichem Nutzen ist, andererseits forderte er von den Vertretern der anderen Fachrichtungen, daß diese in ihren Lehrveranstaltungen, so oft es nur geht, Gebrauch von der Mathematik machen. Die Frage, welchen Platz die Anwendungsgebiete in der Mathematik-Ausbildung einnehmen sollen, spielte für ihn die zentrale Rolle, und er stellte als Ziel heraus, daß zwischen Mathematik und Praxis keine Stoßstelle auftreten darf, sondern ein nahtloser Übergang möglich sein sollte. Man wird auch seiner Überzeugung nach der Bedeutung der Mathematik nicht mehr gerecht, wenn man ihr die Rolle einer Hilfswissenschaft zuschreibt; vielmehr werden in steigendem Maße effektive mathematische Methoden bereitgestellt, von denen entscheidende Impulse für die Mathematik als Produktivkraft ausgehen.

Die wissenschaftliche Arbeit hat bei HEINRICH nie geruht. In einer Veröffentlichung, die noch während seines Aufenthaltes in der Sowjetunion erschien, werden die Untersuchungen zur graphischen Integration mit einem Genauigkeitsvergleich für Halbschrittverfahren wieder aufgenommen, und sehr bald nach seiner Rückkehr in die DDR folgten Arbeiten zu den oben erwähnten Komplexen Rechenschemata und Polynomgleichungen. Danach wandte er sich den Fragen der Lokalisierung und Berechnung von Matrizeneigenwerten zu. Dieses Problem war deshalb von besonderer Aktualität, weil die vorhandenen Kenntnisse nicht ausreichten, um einen zweckmäßigen Einsatz der modernen Rechentechnik zu gewährleisten.

Heinrich gibt dem Verfahren von Voetter, einem Verfahren zur Berechnung des charakteristischen Polynoms einer Matrix, eine Form, die für das Verständnis der Wirkungsweise und für die automatische Rechnung besonders zweckmäßig ist, und weist auf die Beziehung zu dem bekannten Verfahren von Hessenberg hin. In einer Reihe von Arbeiten geht es um die Eingrenzung der Eigenwerte. Die Betragsabschätzung durch eine Norm wird mit der Spektralverschiebung gekoppelt, so daß man eine unendliche Mannigfaltigkeit von Kreisen in der komplexen Ebene erhält. Es erhebt sich die Frage nach dem Kreis mit kleinstem Radius und nach dem Durchschnitt aller Kreise, in dem ja gleichfalls alle Eigenwerte liegen. Bei Verwendung der euklidischen Norm gelingt es Heinrich, diese Frage vollständig zu beantworten; 10 Jahre später löst er das gleiche Problem unter Verwendung der Zeilensummennorm. Die weitere Untersuchung ging in zwei Richtungen. Zur Verschärfung der Abschätzung wird eine weitere einfach zu berechnende Kenngröße herangezogen; man erhält dabei elliptisch geformte Einschließungsgebiete. Außerdem läßt sich durch Ähnlichkeitstransformationen die euklidische Norm verkleinern, ohne daß sich die Eigenwerte ändern; für den Fall, daß zur Transformation eine reelle Diagonalmatrix benutzt wird, gibt Heinrich ein Verfahren zur schrittweisen Minimierung der Norm an. Diese Arbeiten zeigen, welche Bedeutung die zur Abschätzung tatsächlich genutzte Information besitzt. Ähnliche Gedanken spielen bei der Diskussion über ein Konditionsmaß für Matrizen eine Rolle. Auch mit dem sogenannten inversen Eigenwertproblem, das sowohl in der Molekülspektroskopie wie im Flugzeugbau aufgetreten ist, hat sich Heinrich beschäftigt. Er bereicherte die numerische Mathematik noch um viele weitere Anregungen, die er in zahlreichen Vorträgen weitergab oder die unmittelbar Eingang in seine Vorlesungen fanden. Seine wissenschaftlichen Leistungen wurden im Jahre 1964 durch die Berufung zum ordentlichen Mitglied der Deutschen Akademie der Naturforscher (Leopoldina) zu Halle gewürdigt.

Noch in einer zweiten Hinsicht trat Heinrich die Nachfolge von Willers an, nämlich im Jahre 1959 als Herausgeber der „Zeitschrift für Angewandte Mathematik und Mechanik" (ZAMM), also jener Zeitschrift, die sich seit ihrer Gründung im Jahre 1921 weltweite Anerkennung erworben hatte und seit 1947 vom Akademie-Verlag Berlin herausgegeben wird. Die Zeitschrift erlebte unter seiner Leitung einen weiteren Aufschwung; Auflagenhöhe und Umfang mußten mehrmals erhöht werden, und ein international hochgeachtetes Herausgebergremium wurde gewonnen. Heinrich gewann rasch Autorität auch als Herausgeber und hatte sehr bald Kontakte mit hervorragenden Wissenschaftlern aus aller Welt. Viele von ihnen besuchten Dresden und bereicherten durch Fachvorträge das mathematische Leben an der Technischen Hochschule Dresden. Durch den Zeitschriftenaustausch der ZAMM wurde die Palette der sofort zugänglichen Fachzeitschriften entscheidend erweitert. Die „Gesellschaft für Angewandte Mathematik und Mechanik", über deren Jahrestagungen die ZAMM regelmäßig in einem Sonderheft berichtete, wählte Heinrich zunächst in ihren wissenschaftlichen Ausschuß und dann als zweiten Vorsitzenden in den Vorstandsrat. Zwar ergab sich hieraus für seine Tätigkeit manche Anregung, aber der in dieser Zeit auch auf dem Gebiet der Wissenschaftspolitik betriebene Alleinvertretungsanspruch der BRD gegenüber der DDR führte zum Abbruch dieser Verbindungen.

Da sich der Mangel an einem Lehrbuch, das den aktuellen Stand der Grundlagen der numerischen Mathematik wiedergab, immer störender bemerkbar machte, faßte Heinrich den Plan, ein auf drei Bände veranschlagtes Buch „Einführung in die praktische Analysis“ zu schreiben. Hier sollten seine in Lehre und Forschung gesammelten Erfahrungen ihren Niederschlag finden. Das Übermaß an Arbeit, das auf ihm ruhte, hatte jedoch zur Folge, daß davon nur der erste Band erschien. Ein weiteres bedeutsames publizistisches Vorhaben konnte Heinrich dadurch realisieren, daß er eine Reihe von älteren Mitarbeitern als Mitautoren gewann. In kurzer Folge wurden für eine Anzahl von Taschenbüchern für Ingenieure die Mathematik-Teile verfaßt. Er konnte dabei das Ergebnis seiner pädagogischen Erfahrungen einem großen Leserkreis vorstellen und vor allem dafür sorgen, daß den Ingenieuren die Elemente der numerischen Mathematik nahegebracht wurden. Es sei noch ergänzend erwähnt, daß Heinrich viele Jahre für die Zeitschrift „Mathematik in der Schule“ als Redaktionsmitglied tätig war und als einer der beiden Herausgeber der Reihe „Mathematik für Naturwissenschaft und Technik“ des VEB Deutscher Verlag der Wissenschaften Berlin wirkte und bis auf den heutigen Tag wirkt.

Nicht lange konnte sich Heinrich seinen unmittelbaren Aufgaben mit ungeteilten Kräften widmen. Die akademischen Ämter, die er zwischen 1958 und 1968 bekleidete, können hier nur summarisch aufgezählt werden: Fachrichtungsleiter für Mathematik, Abteilungsvorstand, Wahlsenator, Dekan und Prodekan. Er machte sich rasch mit seinem jeweils neuen Pflichtenkreis vertraut und führte die Amtsgeschäfte mit nie erlahmendem Elan. Der von ihm maßgeblich ausgearbeitete, speziell auf die Verhältnisse an der TH Dresden zugeschnittene Mathematik-Studienplan hat sich sehr gut bewährt; viele der hier gesammelten Erfahrungen sind nicht ohne Einfluß auf die späteren Ausbildungspläne geblieben. Es ist ferner wesentlich seiner Initiative zu danken, daß an der TU Dresden das Lehrerstudium in den mathematisch-naturwissenschaftlichen Disziplinen eingeführt wurde, speziell für den zum Abitur führenden Zweig der berufsbildenden Schulen. Bei den in diesem Zusammenhang geführten Diskussionen über das Verhältnis von Fachausbildung und methodischer Ausbildung von Lehrern ging es ihm vor allem um die Erkenntnis, daß eine Erörterung methodischer Fragen völlige Klarheit über die fachlichen Grundlagen voraussetzt. — Heinrich war es auch, der darauf drängte, daß die neuentwickelten mathematischen Grundlagen der Operationsforschung, besonders der Optimierung, in die Ausbildung der Ingenieurökonomen aufgenommen wurden, und der es erreichte, daß der Lehrkörper dieser Fakultät an einer entsprechenden Weiterbildung teilnahm. Nur am Rande sei noch angemerkt, daß er auch an vielen weiteren Stellen wirksam wurde; er hielt z. B. allgemeinbildende Sonntagsvorträge, trug an anderen Stellen seine Vorstellungen über die mathematische Modellbildung vor und wurde von verschiedenen Organen der Volksbildung zu beratender Tätigkeit herangezogen.

Der Beginn der 3. Hochschulreform brachte noch einmal, ein Jahr vor seinem 65. Geburtstag, eine einschneidende Veränderung in Heinrichs Tätigkeit. Er folgte, getragen von einem eindeutigen Vertrauensbeweis seiner Kollegen, dem Ruf als erster Direktor der neugegründeten Sektion Mathematik. Auf der Grundlage eines tiefen persönlichen Verständnisses für die notwendigen gesellschaftlichen und politischen Veränderungen und für das Anliegen der 3. Hochschulreform setzte er sich mit ganzer Kraft für die Neuorganisation und inhaltliche Umgestaltung sowohl des Ausbildungs- und Erziehungsprozesses als auch des Forschungsprozesses ein. Damals entstand der Gedanke zur Bildung von Lehr- und Forschungskollektiven, und es wurde neben drei Wissenschaftsbereichen an der Sektion Mathematik ein Bereich Allgemeine Mathematik gegründet, der der Konzentration derjenigen Kräfte diente, die vorwiegend in der mathematischen Grundausbildung der Ingenieure, Naturwissenschaftler und Ökonomen eingesetzt sind. Unter der Leitung von HEINRICH wurden die wesentlichen Grundlagen für die zukünftige erfolgreiche Entwicklung der Sektion Mathematik gelegt.

Besonders bemerkenswert ist der neue Leitungsstil, den HEINRICH zu dieser Zeit entwickelt. Er erweist sich als unermüdlicher Kämpfer für den Fortschritt, für Überwindung verschiedener Hemmnisse und die Beseitigung mancher längst überholter Traditionen und Vorurteile und bewältigt so die nicht in jeder Phase problemlose Konsolidierung der Sektion. Seine Handlungsweise ist von einer überzeugenden positiven Haltung zur sozialistischen Gesellschaft und zum sozialistischen Staat, der Deutschen Demokratischen Republik, geprägt. HEINRICH ist immer bereit, seine reichen Erfahrungen weiter zu vermitteln, und er hat stets ein offenes Ohr für die Probleme und Sorgen seiner Mitarbeiter und Studenten. Bei ihm prägen sich im Zuge der Entwicklung bemerkenswerte Züge und Eigenschaften eines sozialistischen Leiters und Hochschullehrers aus. So wird HEINRICH in dieser Zeit erneut Vorbild für viele seiner Kollegen und vor allem für die jüngere Generation.

Auch nach seiner Emeritierung zum 31. 8. 1971 ist er der Sektion verbunden geblieben, er widmet sich jetzt wieder hauptsächlich der Ausbildung im Fachstudium Numerische Mathematik. Am 31. 5. 1972 wurde ihm von der Technischen Hochschule Wien die Würde eines Doktors der technischen Wissenschaften ehrenhalber verliehen.

Die hohe Wertschätzung, die HEINRICH von seinen Kollegen, seinen Mitarbeitern aus dem wissenschaftlichen wie dem nichtwissenschaftlichen Bereich und seinen Studenten stets entgegengebracht wird, beruht nicht allein auf den hier geschilderten umfassenden Verdiensten, sondern gleichermaßen auf der Ausstrahlungskraft, die im persönlichen Umgang mit ihm spürbar ist. Er ist vielseitig gebildet und informiert, so daß es kaum ein Gebiet gibt, über das man sich mit ihm nicht mit Gewinn unterhalten könnte. Viele verdanken ihm Hinweise, die für den Fortgang einer Arbeit manchmal aber auch für die ganze Arbeitsrichtung entscheidend gewesen sind. Interessiert verfolgte er die Entwicklung der ihm anvertrauten jungen Wissenschaftler und bemühte sich, jeden so zu fördern, daß er seinen Fähigkeiten und Interessen entsprechend bestmöglich arbeiten konnte. Durch seine Herausgebertätigkeit einem hohen wissenschaftlichen Niveau verpflichtet, achtet er in seiner Umgebung auf die gleichen anspruchsvollen Maßstäbe. In der Bewältigung eines umfangreichen Arbeitspensums ist er allen — auch den wesentlich Jüngeren — stets ein vielbewundertes Vorbild. Die Erklärung ist wohl darin zu suchen, daß HEINRICH von jeher ein begeisterter Hochschullehrer war, dem die Heranbildung eines akademischen Nachwuchses, der den Aufgaben des Berufes in jeder Hinsicht gewachsen ist, immer als erstes am Herzen liegt. Man spürt dies auch besonders, wenn er mit Studenten zu tun hat; der Altersunterschied bedeutet keine Schranke für eine gute gegenseitige Verständigung, und der größere Schwung und der stärkere Optimismus ist oft auf der Seite des Älteren. Daß dabei auch der Humor nicht zu kurz kommt, ist jedem vertraut, der ihn persönlich kennt. Mit seiner Emeritierung ist eine Persönlichkeit aus dem Amt geschieden, die das Wissenschaftsgebiet Mathematik an der TU Dresden in den vergangenen zwei Jahrzehnten entscheidend mitgeformt und sich für dessen Förderung mit dem ganzen Gewicht seiner Autorität eingesetzt hat.

Es wird wohl allen, die ihn in seiner beschwingten, frischen Lebensart kennen, schwerfallen, den Begriff eines Jubilars mit dem Namen HEINRICH in Verbindung zu bringen. Doch der Eintritt ins achte Lebensjahrzehnt ist in jedem Fall ein wichtiger Lebensabschnitt, und Fachkollegen, Schüler und Mitarbeiter können ihm von ganzem Herzen nichts Besseres wünschen, als daß Gesundheit und Schaffenskraft ihm auch weiter erhalten bleiben und daß er im häuslischen Kreis und bei der wissenschaftlichen Arbeit viel Freude erleben möge.

G. OPITZ, W. WINKLER

Sektion Mathematik
der Technischen Universität Dresden

Inhalt

Beiträge zur Numerischen Mathematik
4 (1975), 15–22

Kombination zweier Monotoniesätze von Redheffer und J. Schröder zur numerischen Konstruktion von Schrankenfunktionen

RAINER ANSORGE und HORST KRETH

Herrn Prof. Dr. Dr. h. c. H. HEINRICH zum 70. Geburtstag gewidmet

Es sei B ein offenes zusammenhängendes beschränktes Gebiet des $\mathbb{R}^m$ mit Rand Γ. Es seien u, v zwei Funktionen aus $C^2(B)$. Bei dem Operator

$$T[z] := -\sum a_{jk}(x, z_l)\, z_{jk} + a(x, z, z_l)$$

sei

a) $a(x, v, v_l)$ in v monoton nicht fallend, d. h., für die gewählte Funktion v gelte

$$a(x, v + k, v_l) - a(x, v, v_l) \geqq 0 \quad \text{für alle} \quad k \geqq 0;$$

b) die Matrix

$$\big(a_{jk}(x, v_l)\big) \geqq 0$$

sei positiv semidefinit;

c) zu jeder in sich kompakten Teilmenge $S \subset B$ existiere eine Funktion $c(x) \in C^2(S)$ und eine Konstante C mit

$$\left.\begin{aligned} &\sum a_{jk}(x, v_l)\, c_j c_k > 1, \\ &\sum a_{jk}(x, v_l)\, c_{jk} \geqq -C \end{aligned}\right\} \qquad \text{für } x \in S;$$

d) es existiere eine in einem Intervall $0 < p \leqq p_0$ erklärte positive und isotone Funktion $g(p)$ mit

$$\int_0^{p_0} \frac{dp}{g(p)} = \infty \qquad \text{(Osgoodsche Forderung an } g(p)\text{)}$$

und den Eigenschaften

$$\left.\begin{aligned} &|a_{jk}(x, v_l + r_l) - a_{jk}(x, v_l)| \leqq g(|\operatorname{grad} r|), \\ &|a(x, u, v_l + r_l) - a(x, u, v_l)| \leqq g(|\operatorname{grad} r|) \end{aligned}\right\} \qquad \text{für } r \in C^2(B).$$

Dann gilt folgender Satz von REDHEFFER [1] (vgl. auch [2], S. 306[1])):

[1]) Ein dort die Voraussetzung a) entstellender Druckfehler ist hier korrigiert.

Besitzen die beiden Funktionen u, v überdies die Eigenschaften

$$T[u] \leqq 0, \quad T[v] \geqq 0 \quad \textit{in } B \textit{ und } \quad u \leqq v \quad \textit{auf } \Gamma,$$

so folgt

$$u \leqq v \quad \textit{in } B.$$

Dieser Monotoniesatz kann numerisch in folgender Weise ausgenutzt werden: Gesucht sei eine Lösung der Randwertaufgabe

$$T[z] = 0 \quad \text{in } B,$$
$$R[z] = 0 \quad \text{auf } \Gamma,$$

d. h. eine Lösung der Aufgabe

$$M[z] := \{T[z], R[z]\} = 0, \tag{1}$$

wobei mit T auch M im Sinne des obigen Satzes inversmonoton sei (was z. B. im Fall der ersten Randwertaufgabe erfüllt ist).

Das Problem (1) besitze eine Lösung z. Die Voraussetzungen a) bis d) seien für alle $u \in U$ bzw. für alle $v \in V$ erfüllt. Kann man sich dann Näherungen $\hat{u} \in U$ und/oder $\hat{v} \in V$ für die gesuchte Lösung z (mit $\hat{u} \leqq z \leqq \hat{v}$ auf Γ) verschaffen, so gilt

$$\left.\begin{array}{l} z \leqq \hat{v}, \quad \text{sofern} \quad z \in U, \\ \text{bzw.} \\ \hat{u} \leqq z, \quad \text{sofern} \quad z \in V, \\ \text{und daher die Einschließung} \\ \hat{u} \leqq z \leqq \hat{v} \quad \text{für} \quad z \in U \cap V. \end{array}\right\} \tag{2}$$

Solche Näherungen $\hat{u}, \hat{v}$ kann man z. B. auf folgendem Wege (vgl. [3], S. 18) zu gewinnen suchen. Man wähle zwei Funktionenscharen

$$U_\alpha = \{u(x, \alpha_1, \ldots, \alpha_m)\}, \qquad V_\beta = \{v(x, \beta_1, \ldots, \beta_n)\}$$

mit

$$U_\alpha \subset C^2(B) \quad \text{für alle} \quad \alpha = (\alpha_1, \ldots, \alpha_m) \in A_m \subset \mathbb{R}^m,$$
$$V_\beta \subset C^2(B) \quad \text{für alle} \quad \beta = (\beta_1, \ldots, \beta_n) \in B_n \subset \mathbb{R}^n.$$

Anschließend versuche man die Bestimmung geeigneter Scharparameter α, β aus der Optimierungsaufgabe

$$0 \leqq v(x, \beta) - u(x, \alpha) \leqq \delta \quad \text{für} \quad x \in B \cup \Gamma, \quad \delta \doteq \underset{A_m, B_n}{\text{Min}} \tag{3}$$

mit den (unendlich vielen) Nebenbedingungen

$$(M[u])\,(x, \alpha) \leqq 0 \leqq (M[v])\,(x, \beta) \quad \text{für alle} \quad x \in B \cup \Gamma. \tag{4}$$

Existieren Lösungen $\hat{u} = u(\cdot, \hat{\alpha})$ und $\hat{v} = v(\cdot, \hat{\beta})$ mit

$$\hat{u} \in U \quad \text{bzw.} \quad \hat{v} \in V,$$

so gilt (2), wenn noch $\hat{u} \leqq z \leqq \hat{v}$ auf Γ gesichert ist. Man hat also vor Lösung der Aufgabe (3), (4) sicherzustellen, daß $U_\alpha \cap U \neq \emptyset$ und $V_\beta \cap V \neq \emptyset$ ist; man braucht jedoch nicht notwendig die als zusätzliche Restriktionen aufzufassenden Forderungen $U_\alpha \subset U$, $V_\beta \subset V$ zu erheben, muß dann allerdings a posteriori $\hat{u} \in U \cap U_\alpha$ bzw. $\hat{v} \in V \cap V_\beta$ überprüfen.

Der Aufwand zur Lösung einer Aufgabe der Form (3), (4) ist aber in der Regel erheblich, so daß man (gegebenenfalls unter Inkaufnahme schlechterer Einschließungen (2)) nach einfacher zu handhabenden Verfahren fragen muß.

So kann man z. B. trachten, statt (3), (4) die einseitigenApproximationsaufgaben (vgl. [3], S. 19)

$$-\delta_1 \leqq (M[u])\,(x,\alpha) \leqq 0, \quad \delta_1 \doteq \underset{A_m}{\mathrm{Min}} \tag{5}$$

und

$$0 \leqq (M[v])\,(x,\beta) \leqq \delta_2, \quad \delta_2 \doteq \underset{B_n}{\mathrm{Min}} \tag{6}$$

zu lösen.

Wir setzen im folgenden voraus, daß Minimallösungen $\hat{u}, \hat{v}$ der Aufgaben (5), (6) bezüglich U_α bzw. V_β existieren und ermittelt seien, wobei U_α und V_β zugleich so gewählt seien, daß sowohl jedes Element von U_α als auch jedes Element von V_β den Randbedingungen der gegebenen Aufgabe (1) genügt. (5), (6) reduziert sich dann auf

$$-\delta_1 \leqq (T[u])\,(x,\alpha) \leqq 0, \quad \delta_1 \doteq \mathrm{Min}, \tag{5'}$$

$$0 \leqq (T[v])\,(x,\beta) \leqq \delta_2, \quad \delta_2 \doteq \mathrm{Min}. \tag{6'}$$

Ist

$$T^{\#}[z] = 0 \quad \text{in } B, \qquad R[z] = 0 \quad \text{auf } \Gamma \tag{7}$$

eine andere Randwertaufgabe (bei gleichen Randbedingungen), welche die gleiche Lösung z besitzt wie die Aufgabe (1), erfüllt auch $T^{\#}$ die Voraussetzungen a) bis d) des Redhefferschen Satzes für alle $u \in U^{\#}$ bzw. für alle $v \in V^{\#}$, ist überdies $U_\alpha \cap U^{\#} \neq \emptyset$ bzw. $V_\beta \cap V^{\#} \neq \emptyset$, so liegt es nahe, neben (5'), (6') auch die einseitigen Approximationsaufgaben

$$-\delta_1^{\#} \leqq (T^{\#}[u])\,(x,\alpha) \leqq 0, \quad \delta_1^{\#} \doteq \mathrm{Min} \qquad (\text{bezüglich } U_\alpha), \tag{8}$$

$$0 \leqq (T^{\#}[v])\,(x,\beta) \leqq \delta_2^{\#}, \quad \delta_2^{\#} \doteq \mathrm{Min} \qquad (\text{bezüglich } V_\beta) \tag{9}$$

zu betrachten. Dies ist allerdings nur dann sinnvoll, wenn eine Lösung $\tilde{u} \in U_\alpha \cap U^{\#}$ bzw. $\tilde{v} \in V_\beta \cap V^{\#}$ existiert, auch z die Eigenschaft $z \in U^{\#}$ oder $z \in V^{\#}$ besitzt (so daß (2) mit $\tilde{u}, \tilde{v}$ anstelle von $\hat{u}, \hat{v}$ gilt) und

$$u^* := \sup(\hat{u}, \tilde{u}) \neq \hat{u} \quad \text{bzw.} \quad v^* := \inf(\hat{v}, \tilde{v}) \neq \hat{v} \tag{10}$$

erfüllt ist: u^* bzw. v^* stellen dann verbesserte Schrankenfunktionen dar.

Bemerkung. Unterscheidet sich $T^{\#}$ in gewissem Sinne nicht zu stark von T, so wird man zur Lösung von (8), (9) im wesentlichen mit dem gleichen Rechenprogramm auskommen, das zur Lösung von (5'), (6') benutzt wurde (in der Regel sind nur an wenigen Stellen des Programms geringfügige Modifizierungen vorzunehmen). Unter Umständen ist damit die Schrankenverbesserung (10) numerisch einfacher zu realisieren und nicht viel weniger effektiv als eine Verbesserung durch Erhöhung der Parameteranzahl m oder n.

In [4] beschreibt J. Schröder (dort im wesentlichen zur Klärung von Existenzfragen) bei vorausgesetzter Kenntnis von Funktionen $\hat{u}$, $\hat{v}$ mit

$$T[\hat{u}] \leqq 0 \leqq T[\hat{v}]$$

(wofür wir unsere Lösungen von (5'), (6') verwenden wollen) einen Übergang von T zu einem Operator $T^{\#}$, der unter gewissen Voraussetzungen eine Lösung z der Aufgabe $T^{\#}[z] = 0$, $R[z] = 0$ in dem Intervall $\langle \hat{u}, \hat{v} \rangle$ liefert, die zugleich auch Lösung von (1) ist, und der häufig so konstruiert werden kann, daß er sich im oben skizzierten Sinne hinsichtlich des rechentechnischen Aufwandes nur wenig von T unterscheidet.

Dazu setze man (in unserem Fall für $w \in C^2(B)$)

$$w^{\#} := \sup \{\hat{u}, \inf (w, \hat{v})\}. \tag{11}$$

Schreibt man $(T[w])(x) = T(x, w, w_x, w_{xx})$[1]) mit einer reellwertigen Funktion $T(x, w, p, q)$, so konstruiere man den Operator $T^{\#}$ gemäß

$$(T^{\#}[w])(x) = T^{\#}(x, w, w^{\#}, w_x, w_{xx}) \tag{12}$$

mit einer reellwertigen Funktion $T^{\#}(x, w, y, p, q)$, die die Eigenschaft

$$T^{\#}(x, w, w, p, q) = T(x, w, p, q) \tag{13}$$

besitzen möge. Mit diesem $T^{\#}$ formulieren wir nun die Aufgabe (7), womit wegen (11) sofort klar wird, daß jede in $\langle \hat{u}, \hat{v} \rangle$ liegende Lösung dieser Aufgabe auch Lösung von (1) ist und umgekehrt jede in $\langle \hat{u}, \hat{v} \rangle$ liegende Lösung von (1) auch Lösung von (7) wird. Um $z \in \langle \hat{u}, \hat{v} \rangle$ gewährleisten zu können, fordern wir jetzt $z \in U \cap V$, so daß in (2) wirklich die beidseitige Einschließung gilt. Gemäß dem Umstand, daß in

$$T(x, z, z_x, z_{xx}) = -\sum a_{jk}(x, z_l)\, z_{jk} + a(x, z, z_l)$$

nur das Glied $a(x, z, z_l)$ explizit von z abhängt, spezialisieren wir uns auf den Fall

$$\left.\begin{aligned} &T^{\#}(x, w, y, w_x, w_{xx}) = -\sum a_{jk}(x, w_l)\, w_{jk} + a^{\#}(x, w, y, w_l) \\ &\text{mit}\quad a^{\#}(x, w, w, w_l) = a(x, w, w_l). \end{aligned}\right\} \tag{14}$$

Dann erfüllt auch $T^{\#}$ für alle $u \in U$ und für alle $v \in V$ zunächst die Bedingungen b) und c) des Redhefferschen Satzes sowie den auf die a_{jk} bezogenen Teil der Voraussetzung d). Auch den auf die Funktion a bezogenen Teil der Voraussetzung d) können wir ohne große Einschränkungen der mannigfachen Konstruktionsmöglichkeiten von $a^{\#}$ als durch die $v \in V$ erfüllt ansehen, da diese Voraussetzung im wesentlichen nur die Variable v_l betrifft.

Eine echte Einschränkung der Konstruktionsmöglichkeiten für $a^{\#}$ sowie gegebenenfalls der Funktionen $v \in V$ bedeutet der Wunsch nach Beibehaltung des Erfülltseins der Voraussetzung a).

---

[1]) x steht abkürzend für den Vektor der Ortsvariablen x_i, z_x bzw. z_{xx} für die Matrizen der ersten bzw. zweiten Ableitungen.

Wir nehmen deshalb an: $a^\#$ sei so konstruiert und eine geeignete Menge $V^\# \subset V$ so ausgewählt, daß für alle $v \in V^\#$ und für alle $k \geqq 0$ folgende Bedingung gelte:

$$a^\#(x, v + k, (v + k)^\#, v_l) - a^\#(x, v, v^\#, v_l) \geqq 0. \tag{15}$$

Wir können also $U^\# = U$ setzen und wählen $V^\# \subset V$ geeignet. (16)

Auch durch die Forderung (15) wird jedoch die Konstruktion von $a^\#$ nicht eindeutig festgelegt, so daß hier für das Folgende noch Spielraum verbleibt.

Bemerkung. In [4] wird unter den dort genannten Voraussetzungen sichergestellt, daß jede Lösung von (7) mit dem gemäß (13) konstruierten Operator $T^\#$ (sofern solche Lösungen existieren) in dem Intervall $\langle \hat{u}, \hat{v} \rangle$ liegt. Da jedoch bereits die Existenz einer Lösung $z \in U \cap V$ (und damit $z \in \langle \hat{u}, \hat{v} \rangle$) der Aufgabe (1) vorausgesetzt wurde, können die Monotonievoraussetzungen aus [4] hier außer Betracht bleiben.

Es seien nun Minimallösungen $\tilde{u}$ (bezüglich U_α) bzw. $\tilde{v}$ (bezüglich V_β) der Aufgabe (8) bzw. (9) für den gemäß (14) konstruierten Operator $T^\#$ gefunden. Dann gilt (unter Beachtung des Umstandes, daß die Randbedingung durch alle Elemente von U_α und V_β erfüllt werden) zunächst wenigstens folgender

Satz. *Es ist* $\delta_i{}^\# \leqq \delta_i$ $(i = 1, 2)$.

Beweis. Auf Grund der Extremaleigenschaft von $\tilde{u}$ gilt

$$\begin{aligned} -\delta_1{}^\# &= \inf_{x \in B} T^\#\big(x, \tilde{u}(x), \tilde{u}^\#(x), \tilde{u}_x(x), \tilde{u}_{xx}(x)\big) \\ &\geqq \inf_{x \in B} T^\#\big(x, \hat{u}(x), \hat{u}^\#(x), \hat{u}_x(x), \hat{u}_{xx}(x)\big). \end{aligned}$$

Nun ist $\hat{u}^\# = \sup\{\hat{u}, \inf(\hat{u}, \hat{v})\} = \sup\{\hat{u}, \hat{u}\} = \hat{u}$, so daß aus

$$\begin{aligned} \inf_{x \in B} T^\#\big(x, \hat{u}(x), \hat{u}^\#(x), \hat{u}_x(x), \hat{u}_{xx}(x)\big) &= \inf_{x \in B} T^\#(x, \hat{u}, \hat{u}, \hat{u}_x, \hat{u}_{xx}) \\ &= \inf_{x \in B} T(x, \hat{u}, \hat{u}_x, \hat{u}_{xx}) = -\delta_1 \end{aligned}$$

folgt, daß $\delta_1{}^\# \leqq \delta_1$ ist. Analog erhält man $\delta_2{}^\# \leqq \delta_2$.

Es kann allerdings (im Sinne der natürlichen Halbordnung in $C^2(B)$) nicht zugleich

$$\tilde{u} \geqq \hat{u} \quad \text{und} \quad \tilde{u} \neq \hat{u} \quad (\text{bzw. } \tilde{v} \leqq \hat{v} \text{ und } \tilde{v} \neq \hat{v})$$

erwartet werden; denn gilt $\tilde{v} \leqq \hat{v}$, d. h. $\hat{u} \leqq z \leqq \tilde{v} \leqq \hat{v}$, so folgt

$$\tilde{v}^\# = \sup\{\hat{u}, \inf(\tilde{v}, \hat{v})\} = \sup\{\hat{u}, \tilde{v}\} = \tilde{v}$$

und damit unter Benutzung der Extremaleigenschaft von $\hat{v}$

$$\begin{aligned} \delta_2{}^\# &= \sup_{x \in B} T^\#\big(x, \tilde{v}(x), \tilde{v}^\#(x), \tilde{v}_x(x), \tilde{v}_{xx}(x)\big) \\ &= \sup_{x \in B} T^\#(x, \tilde{v}, \tilde{v}, \tilde{v}_x, \tilde{v}_{xx}) = \sup_{x \in B} T(x, \tilde{v}, \tilde{v}_x, \tilde{v}_{xx}) \\ &\geqq \sup_{x \in B} T(x, \hat{v}, \hat{v}_x, \hat{v}_{xx}) = \delta_2 \quad (\text{analog } \delta_1{}^\# \geqq \delta_1). \end{aligned} \tag{17}$$

Also würde in diesem Fall in Verbindung mit obigem Satz

$$\delta_i^{\#} = \delta_i \quad (i = 1, 2)$$

und daher gemäß (17)

$$\sup_{x \in B} T(x, \tilde{v}, \tilde{v}_x, \tilde{v}_{xx}) = \sup_{x \in B} T(x, \hat{v}, \hat{v}_x, \hat{v}_{xx}) \quad \text{(analog für } \tilde{u}, \hat{u})$$

resultieren; dies aber bedeutet mindestens bei Eindeutigkeit der Minimallösungen der Aufgaben (5′), (6′):

$$\tilde{u} = \hat{u}, \qquad \tilde{v} = \hat{v}.$$

Andererseits liefern die soeben angestellten Überlegungen, daß $\delta_2^{\#} < \delta_2$ auch wirklich $\tilde{v} \neq \hat{v}$ (analog $\tilde{u} \neq \hat{u}$ im Fall $\delta_1^{\#} < \delta_1$) nach sich zieht (womit allerdings noch nicht (10) gewährleistet ist).

Es lassen sich jedoch Beispiele angeben, für die die Defekte $\delta_i^{\#}$ gegenüber den Defekten δ_i ($i = 1, 2$) bei geeigneter Konstruktion von $T^{\#}$ echt abnehmen und zugleich (10) realisiert wird:

Beispiel. Gegeben sei die nichtlineare Randwertaufgabe (vgl. [3], S. 19)

$$\left.\begin{aligned} T[z] &= -\Delta z + z^2 = 0 \quad \text{in} \quad B = \left\{(x_1, x_2, x_3);\ \sum_{i=1}^{3} x_i^2 = r^2 < 1\right\}, \\ R[z] &= z - 1 = 0 \quad \text{auf} \quad \Gamma = \left\{(x_1, x_2, x_3);\ \sum_{i=1}^{3} x_i^2 = 1\right\} \end{aligned}\right\} \tag{18}$$

(offenbar handelt es sich um ein kugelsymmetrisches Problem, d. h., es ist $z = z(r)$, $\Delta z = z'' + \frac{2}{r} z'$).

Wegen $a(x, v, v_l) = v^2$ sind die Voraussetzungen a) bis d) des Redhefferschen Satzes erfüllt für

$$U = C^2(B) \cap C^0(\bar{B}), \qquad V = \{v \in U;\ v \geqq 0\}.$$

Die Aufgabe (18) besitzt eine überall positive Lösung z, wodurch $z \in U \cap V$ gesichert ist.

In [3] wird $m = n = 2$, $A_m = B_n = \mathbb{R}^2$,

$$U_\alpha = \{u(r) = 1 + (1 - r^2)(\alpha_1 + \alpha_2 r^2);\ (\alpha_1, \alpha_2) \in \mathbb{R}^2\},$$

$$V_\beta = \{v(r) = 1 + (1 - r^2)(\beta_1 + \beta_2 r^2);\ (\beta_1, \beta_2) \in \mathbb{R}^2\}$$

gewählt. Die Funktionenmengen U_α und V_β sind hier also identisch; ihre Elemente erfüllen offenbar die gegebene Randbedingung des Problems (18). Es ist

$$U_\alpha \cap U \neq \emptyset \quad (\text{wegen} \quad U_\alpha \subset U)$$

und

$$V_\beta \cap V \neq \emptyset \quad (\text{jedoch nicht } V_\beta \subset V!).$$

Man erhält die Lösung $\hat{u}$ der Aufgabe (5′) für

$$\alpha_1 = -0.13691, \quad \alpha_2 = -0.01275.$$

Die Lösung $\hat{v}$ der Aufgabe (6′) wird durch

$$\beta_1 = -0.13545, \quad \beta_2 = -0.01263$$

charakterisiert. Man prüft sofort die wichtige Eigenschaft $\hat{v} \in V$ nach, so daß in der Tat die Einschließung

$$\hat{u} \leqq z \leqq \hat{v}$$

gewährleistet ist. Speziell erhält man für den kleinsten Wert der isotonen Funktion $z(r)$ $(0 \leqq r \leqq 1)$

$$0.86309 \leqq z(0) \leqq 0.86455. \tag{19}$$

Bei $r = 0$ nimmt die Streifenbreite $\hat{v}(r) - \hat{u}(r)$ übrigens ihr Maximum an:

$$\hat{v}(0) - \hat{u}(0) = 0.00146.$$

Der Defekt der Oberfunktion, d. h. $-\Delta\hat{v} + \hat{v}^2$, verschwindet bei $r \approx 0.7$.

Wir wählen nun z. B.

$$(T^{\#}[w])(x) = -\Delta w + a^{\#}(r, w, w^{\#}, w_l) \tag{20}$$

mit

$$a^{\#}(r, w, y, p) = a^{\#}(r, w, y) = \begin{cases} w^2 & \text{für } 0 \leqq r \leqq k_1 < 0.7, \\ \left.\begin{array}{ll} w^2 - k_2(r - k_1)^2 & \text{für } w > y \\ w^2 & \text{für } w \leqq y \end{array}\right\} & \text{für } k_1 < r \leqq 1, \end{cases} \tag{21}$$

wobei die Konstanten $k_1 \in \langle 0, 0.7)$ und $k_2 > 0$ noch verfügbar bleiben. Offenbar gilt daher wegen $z^{\#} = z$ in der Tat für obige Lösung z der Aufgabe (1)

$$a^{\#}(r, z, z^{\#}) = z^2 = a(z).$$

Die Monotoniebedingung a) des Redhefferschen Satzes (bezogen auf $a^{\#}$) ist offenbar erfüllt für alle

$$v \in V^{\#} := \{v \in V;\, v(r) > \hat{v}(r) \text{ für } k_1 < r < 1\}.$$

Wir lösen nun die Aufgabe

$$0 \leqq (T^{\#}[v])(x, \beta) \leqq \delta_2{}^{\#} = \delta_2{}^{\#}(k_1, k_2) \doteq \text{Min} \tag{9′}$$

bezüglich $v \in V_\beta$, wobei k_1 und k_2 noch geeignet wählbar sind. Beispielsweise erhält man für $k_1 = 0.6$, $k_2 = 0.10669$ die Lösung $\tilde{v}$ der Aufgabe (9′) für die Parameterwerte

$$\beta_1 = -0.13575, \qquad \beta_2 = -0.01180.$$

Tatsächlich ist

$$\delta_2{}^{\#} = 0.00324 < \delta_2 = 0.01053,$$

und man verifiziert sofort $\tilde{v} \in V^{\#}$. Überdies ist $z \in U^{\#}$ wegen $U^{\#} = U$.

Damit ist in der Tat

$$v^* := \inf(\tilde{v}, \hat{v}) = \begin{cases} \tilde{v}(r) & \text{für} \quad 0 \leqq r \leqq 0.6, \\ \hat{v}(r) & \text{für} \quad 0.6 < r \leqq 1 \end{cases}$$

eine bessere Oberfunktion als $\hat{v}$. Speziell folgt

$$\tilde{v}(0) = 0.86425,$$

womit sich auch ohne Verbesserung der Unterfunktion[1]) bereits eine Verringerung der maximalen Streifenbreite ($\tilde{v}(0) - \hat{u}(0) = 0.00116$) um ca. 20% ergibt.

Literatur

[1] Redheffer, R. M., An extension of certain maximum principles, Mh. Math. Phys. **66** (1962), 32–42.
[2] Collatz, L., Funktionalanalysis und numerische Mathematik, Springer-Verlag, Berlin—Göttingen—Heidelberg 1964.
[3] Collatz, L., und W. Krabs, Approximationstheorie, Teubner, Stuttgart 1973.
[4] Schröder, J., Einschließungsaussagen bei Differentialgleichungen, in: Numerische Lösung nichtlinearer partieller Differential- und Integrodifferentialgleichungen, Lecture Notes in Mathematics, Vol. 267, hrsg. von R. Ansorge und W. Törnig, Springer-Verlag, Berlin—Heidelberg—New York 1972, S. 23–49.

Manuskripteingang: 6. 5. 1974

VERFASSER:
Prof. Dr. Rainer Ansorge und Dr. Horst Kreth, Institut für Angewandte Mathematik der Universität Hamburg

[1]) Eine Verbesserung der Unterfunktion ist in diesem Beispiel ohne zusätzliche Modifikation von $T^{\#}$ und $V^{\#}$ wegen $z \notin V^{\#}$ nicht zu gewährleisten.

Beiträge zur Numerischen Mathematik
4 (1975), 23–32

Einschließung der Nullstellen von Intervallpolynomen

Karl-Heinz Bachmann

Herrn Prof. Dr. Dr. h. c. H. Heinrich zum 70. Geburtstag gewidmet

1. Einführung

Als reelles Intervallpolynom wird eine Menge von Polynomen

$$F(x) = \left\{ f(x) = \sum_{i=0}^{n} a_i x^i : |a_i - c_i| \leqq e_i \right\} \tag{1}$$

mit gegebenen $a_i, c_i, e_i \in R$ und $e_i \geqq 0$ $(i = 0, \ldots, n)$ bezeichnet. Dabei sei R der Raum der reellen Zahlen. Entsprechend kann ein komplexes Intervallpolynom mit $a_i, c_i \in K$ (Raum der komplexen Zahlen) definiert werden. Die Nullstellenmenge eines Intervallpolynoms ist

$$N = \{x_i : f(x_i) = 0 \wedge f(x) \in F(x)\}. \tag{2}$$

Hat ein beliebiges $f(x)$ aus $F(x)$ keine mehrfachen Nullstellen, so zerfällt die Menge N für genügend kleine e_i in n disjunkte Teilmengen. Häufig tritt jedoch auch der Fall auf, daß N in weniger disjunkte Teilmengen zerfällt. Ein Beispiel ist

$$F(x) = \{a_2 x^2 + a_1 x + a_0 : a_2 = 1 \wedge |a_1 + 7| \leqq 0.5 \wedge |a_0 - 12| \leqq 0.5\}. \tag{3}$$

Das in $F(x)$ enthaltene mittlere Polynom

$$f_0(x) = x^2 - 7x + 12 \tag{4}$$

hat die Nullstellen $x_{01} = 3$ und $x_{02} = 4$. Die vier „Eckpolynome"

$$\left.\begin{aligned} f_1(x) &= x^2 - 7.5x + 12.5,\\ f_2(x) &= x^2 - 7.5x + 11.5,\\ f_3(x) &= x^2 - 6.5x + 12.5,\\ f_4(x) &= x^2 - 6.5x + 11.5 \end{aligned}\right\} \tag{5}$$

haben die Nullstellen

$$\left.\begin{aligned} x_{11} &= 2.5, & x_{12} &= 5,\\ x_{21} &= 2.1492, & x_{22} &= 5.3508,\\ x_{31} &= 3.25 + 1.3919i, & x_{32} &= 3.25 - 1.3919i,\\ x_{41} &= 3.25 + 0.9682i, & x_{42} &= 3.25 - 0.9682i. \end{aligned}\right\} \tag{6}$$

Die Menge N besteht hier aus dem reellen Intervall (2.1492, 5.3508) und dem Gebiet der komplexen x-Ebene, das durch $\Re(x) \geqq 3.25$ und $3.391 < |x| < 3.536$ bestimmt, d. h. durch zwei Kreisbögen und eine Gerade begrenzt ist. Es gibt im Intervallpolynom (3) auch Polynome mit Doppelwurzeln, z. B. $(x - 3.4)^2 = x^2 - 6.8x + 11.56$.

Betrachtet man als zweites Beispiel das Intervallpolynom

$$F(x) = \{a_2x^2 + a_1x + a_0 : a_2 = 1 \wedge |a_1 + 7| \leqq 0.05 \wedge |a_0 - 12| \leqq 0.05\}, \tag{7}$$

so zerfällt N in zwei reelle Intervalle

$$N = (2.835, 3.315) \cup (3.635, 4.215). \tag{8}$$

Bei der numerischen Lösung algebraischer Gleichungen ergibt sich das Problem, die Fehler gegebener Näherungen für die Wurzeln abzuschätzen, wenn die Koeffizienten des Polynoms fehlerbehaftet sind. Diese Aufgabe kann auch so formuliert werden, die Nullstellenmenge eines Intervallpolynoms in möglichst kleine Kreise mit gegebenen Mittelpunkten einzuschließen. Offensichtlich ist es nicht in jedem Fall möglich, für die Nullstellenmenge eines Intervallpolynoms n-ten Grades n disjunkte Einschließungskreise anzugeben. Wird für das Intervallpolynom (3) als gemeinsame Näherung für beide Nullstellen 3.5 gewählt, so schließt ein Kreis um 3.5 mit dem Radius 1.8508 das Gebiet N ein. Der kleinste N einschließende Kreis hat den Mittelpunkt 3.75 und den Radius 1.6008. Für das Intervallpolynom (7) lassen sich dagegen zwei disjunkte Einschließungskreise für die Nullstellenmenge angeben, und zwar die Kreise $|x - 3| < 0.315$ und $|x - 4| < 0.365$. Da hier nur reelle Wurzeln auftreten, kann die Abschätzung auf die betreffenden reellen Intervalle verschärft werden.

2. Eine Methode zur Einschließung der Nullstellenmenge

Es sei $f(x)$ ein Polynom n-ten Grades mit den Nullstellen $x_1(f), \ldots, x_n(f)$, das in einem Intervallpolynom $F(x)$ enthalten ist. Das Polynom

$$g(x) = f(x) \Big/ \prod_{i=k+1}^{n} \left(x - x_i(f)\right) \tag{9}$$

hat die Nullstellen $x_1(f), \ldots, x_k(f)$. $x = 0$ kann als gemeinsame Näherung für diese Nullstellen betrachtet werden, wenn man annimmt, daß die $|x_1(f)|, \ldots, |x_k(f)|$ wesentlich kleiner als jeder der Beträge $|x_{k+1}(f)|, \ldots, |x_n(f)|$ sind. Sind weiter Näherungen y_i für die $x_i(f)$ $(i = k + 1, \ldots, n)$ bekannt und gilt

$$|x_i(f) - y_i| \leqq R_i \quad \text{für alle} \quad f(x) \in F(x) \quad \text{und} \quad i = k + 1, \ldots, n, \tag{10}$$

so kann für die Menge von Polynomen

$$G = \left\{g(x) : g(x) = f(x) \Big/ \prod_{i=k+1}^{n} \left(x - x_i(f)\right) \text{ für alle } f(x) \in F(x)\right\} \tag{11}$$

ein Intervallpolynom

$$\bar{G}(x) = \Big\{g(x) : g(x) \text{ ist Resultat eines Divisionsalgorithmus}$$
$$f(x)\Big/\prod_{i=k+1}^{n}(x - Y_i), \text{ wobei } Y_i = \{y : |y - y_i| \leqq R_i\}\Big\} \tag{12}$$

z. B. mit Hilfsmitteln der Intervallarithmetik bestimmt werden, das G als Teilmenge enthält. Jede Schranke für die k Nullstellen enthaltende Nullstellenmenge von $\bar{G}(x)$ ist dann auch Schranke für alle k Nullstellen jedes $g(x)$ in G, d. h. für die Nullstellen $x_1(f), \ldots, x_k(f)$ jedes $f(x)$ in $F(x)$. Eine solche Schranke ist z. B. die Schranke von FUJIWARA [2]

$$S = 2 \cdot \max_{i<k} \sqrt[k-i]{M_i/m_k}, \tag{13}$$

wobei die M_i obere Schranken für die Beträge der i-ten Koeffizienten aller $g(x)$ in $\bar{G}(x)$ sind und m_k eine untere Schranke für den k-ten Koeffizienten aller $g(x)$ in $\bar{G}(x)$ ist. Wird z. B. das Intervallpolynom (3) um $x = 3.5$ entwickelt, so ergibt sich

$$F(x) = \{b_2(x - 3.5)^2 + b_1(x - 3.5) + b_0 :$$
$$b_2 = 1 \wedge |b_1| \leqq 0.5 \wedge |b_0 + 0.25| \leqq 2.25\}. \tag{14}$$

Daraus folgt

$$m_2 = 1, \quad M_1 = 0.5, \quad M_2 = 2.5, \quad S = \sqrt{10} = 3.162. \tag{15}$$

Der Radius 1.85 des oben angegebenen Einschließungskreises mit dem Mittelpunkt 3.5 wird also um etwa 70% überschätzt.

Sind Einschließungen (10) für die Nullstellen $x_{k+1}(f), \ldots, x_n(f)$ nicht bekannt, so kann dennoch ein auf der Division (9) aufbauendes Iterationsverfahren benutzt werden. Näherungen für die Nullstellen seien wie oben mit $y_1, \ldots, y_n$ bezeichnet. Weiter seien die Näherungen in s Klassen $C_1, \ldots, C_s$ zusammengefaßt, die so in s disjunkte Kreise eingeschlossen werden können, daß in diesen jeweils k_j $(j = 1, \ldots, s)$ Nullstellen jedes $f(x)$ aus $F(x)$ enthalten sind. Durch einen Divisionsalgorithmus werden nun s Polynome $g_1(x; \bar{y}_1), \ldots, g_s(x; \bar{y}_s)$ derartig bestimmt, daß

$$g_j(x; \bar{y}_j) \cdot \prod_{r \notin C_j}(x - y_r) = f(x) + r_j(x; \bar{y}_j) \tag{16}$$

gilt, wobei $\bar{y}_j$ eine Abkürzung für die Menge $\{y_r\}_{r \notin C_j}$ ist und $r_j(x; \bar{y}_j)$ ein Restpolynom bezeichnet. Stimmen die y_r $(r \notin C_j)$ mit $n - k_j$ Nullstellen von $f(x)$ überein, so soll dieses Restpolynom identisch verschwinden. Bezeichnet man die Nullstellen von $g_j(x; \bar{y}_j)$ mit $z_1^{(j)}, \ldots, z_{k_j}^{(j)}$, so wird durch die s Divisionen nach (16) und anschließende Nullstellenbestimmungen eine Transformation

$$z = Ty \tag{17}$$

definiert, die die Näherungen y_i $(i = 1, \ldots, n)$ in Näherungen z_i transformiert. Dabei entstehen die z_i aus den $z_r^{(j)}$ $(r = 1, \ldots, k_j;\ j = 1, \ldots, s)$ durch passende

Numerierung. Betrachtet man die Produktpolynome

$$p(x; y) = \prod_{i=1}^{n} (x - y_i) \quad \text{und} \quad p(x; z) = \prod_{i=1}^{n} (x - z_i), \tag{18}$$

so kann auch die Transformation

$$p(x; y) = T' p(x; z) \tag{19}$$

zugrunde gelegt werden, bei der nicht auf eine Numerierung der Wurzeln zu achten ist.

Wenn das Divisionsverfahren so beschaffen ist, daß die Koeffizienten der g_j stetig von den y_r in $\bar{y}_j$ abhängen, so hängen die z_r stetig von den y_r ab. Wird nun ein konvexes Gebiet des n-dimensionalen komplexen Raumes K^n durch T in sich abgebildet, so liegt nach dem Fixpunktsatz von Brouwer (vgl. [1]) ein Fixpunkt der Transformation T in diesem Gebiet. Unter der obigen Voraussetzung über die Restpolynome r_j ist der Punkt $\big(x_1(f), \ldots, x_n(f)\big)$ des K^n ein Fixpunkt der Transformation T. Kann nun gezeigt werden, daß T nur einen Fixpunkt hat (abgesehen von Umnumerierungen der Nullstellen), so enthält das betrachtete Gebiet und daher auch sein Bildgebiet diesen Fixpunkt.

Die Feststellung, ob T ein Gebiet in sich abbildet, kann mit Hilfsmitteln der Fehlerrechnung (Intervallarithmetik) erfolgen. Dazu werden s Näherungen $w_1, \ldots, w_s$ als Repräsentanten der Klassen $C_1, \ldots, C_s$ fest vorgegeben. Im allgemeinen sind es Näherungen, die mit einem numerischen Verfahren für die Nullstellen eines $f(x)$ aus $F(x)$ gefunden werden. Weiter werden Radien R_j für Kreise $|y - w_j| \leqq R_j$ $(j = 1, \ldots, s)$ geschätzt. Durch einen Divisionsalgorithmus werden dann Intervallpolynome $G_j(x)$ bestimmt, die so beschaffen sind, daß die $g_j(x; \bar{y}_j)$ für alle $\bar{y}_j$, die mit y aus den Kreisen $|y - w_j| \leqq R_j$ bildbar sind, und für alle $f(x) \in F(x)$ in $G_j(x)$ enthalten sind. Mit anderen Worten ist hiernach die Division mit intervallarithmetischen Hilfsmitteln so auszuführen, daß $F(x)$ evtl. mehrfach durch lineare Intervallpolynome $x - Y_i$ dividiert wird, wobei $x - Y_i$ die Menge aller $x - y$ mit $|y - w_i| \leqq R_i$ ist. Für jedes $j = 1, \ldots, s$ wird nun eine Schranke S_j bestimmt:

$$|z_i^{(j)} - w_j| \leqq S_j, \tag{20}$$

wobei $g(z_i^{(j)}) = 0$ ist für alle $g(x) \in G_j(x)$. Gilt dann $S_j \leqq R_j$ für alle $j = 1, \ldots, s$, so ist die Voraussetzung des Brouwerschen Fixpunktsatzes erfüllt. Die S_j sind dann obere Schranken für die Abweichungen von jeweils k_j Nullstellen eines beliebigen $f(x) \in F(x)$ von der Näherung w_j, dabei ist k_j der Grad von $G_j(x)$, d. h. die Anzahl der jeweils gemeinsam eingeschlossenen Nullstellen.

3. Algorithmus zur Einschließung

Zur Ausführung der Division nach (16) für alle $f(x) \in F(x)$ wird zunächst $F(x)$ um die Stelle w_j entwickelt, das Resultat der Entwicklung sei als $F_j(t)$ mit $t = x - w_j$ bezeichnet. Anschließend wird $F_j(t)$ durch Linearfaktoren $x - Y_i = t - (Y_i - w_j)$ dividiert $(i \neq j)$. Da vorausgesetzt werden kann, daß die Abstände $Y_i - w_j$ groß