Colin W. Cryer

Numerik Partieller Differentialgleichungen II

Sommersemester[1] 1995

Westfälische
Wilhelms-Universität
Münster

[1] Letzte Änderung am 2. Oktober 1995

Inhaltsverzeichnis

Kapitel 1

Elliptische Gleichungen

1.1 Definition elliptischer Gleichungen

1.1.1 Skalare Gleichungen

Definition 1.1 *Sei*

$$L = \sum_{|\alpha| \leq k} a_\alpha(\mathbf{x}) \partial^\alpha$$

ein Differentialoperator für $\mathbf{x} \in \Omega \subset \mathbb{R}^n$. *L heißt elliptisch zu* $\mathbf{x}$, *falls für alle* $\xi \in \mathbb{R}^n$, $\xi \neq 0$ *gilt*

$$\sigma_x(L, \xi) := \sum_{|\alpha| = k} a_\alpha(\mathbf{x}) \xi^\alpha \neq 0 \; .$$

Beispiel 1.1 Die Laplace-Gleichung in $\mathbb{R}^n$ ist elliptisch, da

$$\begin{aligned} Lu &= \sum_{i=1}^{n} \frac{\partial^2 u}{\partial x_i^2} \; , \\ \sigma_x(L, \xi) &= \sum_{i=1}^{n} \xi_i^2 \; . \end{aligned}$$

Beispiel 1.2 Die biharmonische Gleichung in $\mathbb{R}^2$,

$$Lu = \frac{\partial^4 u}{\partial x_1^4} + 2 \, \frac{\partial^4 u}{\partial x_1^2 \partial x_2^2} + \frac{\partial^4 u}{\partial x_2^4} \; ,$$

ist elliptisch, da

$$\begin{aligned} \sigma_x(L, \xi) &= \xi_1^4 + 2\xi_1^2 \xi_2^2 + \xi_2^4 \\ &= (\xi_1^2 + \xi_2^2)^2 \; . \end{aligned}$$

Beispiel 1.3 Die Minimal-Oberflächengleichung

$$(1+u_y^2)u_{xx} - 2u_x u_y u_{xy} + (1+u_x^2)u_{yy} = 0$$

ist elliptisch, da

$$\begin{aligned}\sigma_x(L,\xi) &= (1+u_y^2)\xi_1^2 - 2u_x u_y \xi_1\xi_2 + (1+u_x^2)\xi_2^2 \,, \\ &= \xi_1^2 + \xi_2^2 + (\xi_1 u_y - \xi_2 u_x)^2\end{aligned}$$

Es ist manchmal wichtig, die Elliptizitätsbedingung genauer zu klassifizieren. Für Differentialgleichungen der zweiten Ordnung ist die folgende Klassifikation nützlich:

Definition 1.2 *Sei*

$$L = \sum_{|\alpha|\le 2} a_\alpha(\mathbf{x})\partial^\alpha$$

ein Differentialoperator zweiter Ordnung für $\mathbf{x} \in \Omega \subset \mathbb{R}^n$. L heißt gleichmäßig elliptisch auf Ω (uniformly elliptic), falls eine Konstante $a > 0$ existiert mit

$$a \sum_{i=1}^{n} \xi_i^2 \le \sum_{|\alpha|=2} a_\alpha(\mathbf{x})\xi^\alpha \le \frac{1}{a} \sum_{i=1}^{n} \xi_i^2$$

für alle $\mathbf{x} \in \Omega$.

Beispiel 1.4 Die Gleichung

$$x u_{xx} + 2u_x = 0$$

ist nicht gleichmäßig elliptisch auf $\Omega = (0,1) \subset \mathbb{R}^1$

1.1.2 Gleichungssysteme

Systeme von Differentialgleichungen treten oft auf. Sie werden allerdings nur selten in der Literatur behandelt. Insbesondere werden Systeme höherer Ordnung kaum in Lehrbüchern behandelt. Einige der wenigen Ausnahmen sind die Bücher von Miranda [1970, S. 275], Courant und Hilbert [1962, S. 577] und Hörmander [1963, S. 267]. Sei

$$\begin{aligned}\mathbf{u}(\mathbf{x}) &= (u_1(\mathbf{x}), \ldots, u_N(\mathbf{x}))^T \,, \\ \mathbf{f}(\mathbf{x}) &= (f_1(\mathbf{x}), \ldots, f_N(\mathbf{x}))^T\end{aligned}$$

für $\mathbf{x} \in \Omega \subset \mathbb{R}^n$. Sei $\mathcal{M}_{ij}$ ein Differentialoperator der Ordnung α_{ij}.

Sei

$$s_i \in \mathbb{Z} \;, \quad t_i \in \mathbb{Z} \;, \quad 1 \leq i \leq N$$

mit

$$\begin{aligned} \alpha_{ij} &\leq s_i + t_j \;, \\ \mathcal{M}_{ij} &\equiv 0 \text{ falls } \quad s_i + t_j < 0 \;. \end{aligned}$$

Sei $M_{ij}(\mathbf{x}, \xi)$ das charakteristische Polynom von $\mathcal{M}_{ij}$ und

$$\hat{M}_{ij}(\mathbf{x}, \xi) = \begin{cases} M_{ij}(\mathbf{x}, \xi) & , \text{ falls } \quad \alpha_{ij} = s_i + t_j \\ 0 & , \text{ falls } \quad \alpha_{ij} < s_i + t_j \end{cases}$$

Das System

$$\sum_{j=1}^{N} \mathcal{M}_{ij}(\mathbf{x}) u_j = f_i \;, \quad 1 \leq i \leq N$$

heißt *elliptisch im Sinne von Douglis und Nirenberg*, falls es möglich ist, s_i und t_j so zu wählen, daß

$$\hat{a}(\mathbf{x}, \xi) := \det(\hat{M}_{ij}(\mathbf{x}, \xi)) \neq 0 \;, \quad \mathbf{x} \in \Omega \;, \quad \xi \in \mathbb{R}^n \;.$$

Sei

$$\begin{aligned} a(\mathbf{x}, \xi) &:= \det(M_{ij}(\mathbf{x}, \xi)) \;, \\ r &= \text{Grad } a \;, \\ R &= \text{Der maximale Grad der N! Terme von } a \\ & \text{ bei der Entwicklung mit der Leibniz-Regel} \end{aligned}$$

Es ist von Volevich bewiesen worden (Agmon, Douglis, Nirenberg [1964, S. 39]), daß die folgenden Bedingungen für die Elliptizität notwendig und hinreichend sind:

1. $r = R$,

2. $\tilde{a}(\mathbf{x}, \xi) \neq 0$ für alle $\xi \neq 0$, wo bei $\tilde{a}$ der Hauptteil von a, d.h. der Teil vom Grade r, ist.

Beispiel 1.5

$$\begin{aligned} u_x - v_{yy} &= f_1 \;, \\ u_{yy} + u_y + v_{xxx} + v_{xy} &= f_2 \end{aligned}$$

Man setze

$$s_1 = 0\,, \quad s_2 = 1\,, \quad t_1 = 1\,, \quad t_2 = 2\,.$$

Es folgt:

$$a = \tilde{a} = \hat{a} = \begin{vmatrix} \xi & -\eta^2 \\ \eta^2 & \xi^3 \end{vmatrix} = \xi^4 + \eta^4\,, \quad r = 4\,, \quad R = 4\,.$$

Beispiel 1.6 [Cauchy-Riemann-Gleichungen]

$$\begin{aligned} u_x - v_y &= 0 \\ u_y + v_x &= 0 \end{aligned}$$

Man setze $s_1 = s_2 = 1$, $t_1 = t_2 = 0$. Es folgt:

$$\hat{a} = \begin{vmatrix} \xi & -\eta \\ \eta & \xi \end{vmatrix} = \xi^2 + \eta^2\,.$$

$$\begin{aligned} \tilde{a} = a &= \hat{a}\,, \\ r = R &= 2\,. \end{aligned}$$

Beispiel 1.7

$$u_{xx} + u_{yy} = 0\,.$$

Setze

$$u_1 := u_x\,, \quad u_2 := u_y\,.$$

Es gilt:

$$\begin{aligned} \frac{\partial u_1}{\partial x_1} + \frac{\partial u_2}{\partial x_2} &= 0\,, \\ \frac{\partial u}{\partial x_1} - u_1 &= 0\,, \\ \frac{\partial u}{\partial x_2} - u_2 &= 0\,. \end{aligned}$$

Man setze

$$\begin{aligned} s_0 = 0\,, \quad s_1 = s_2 &= -1\,, \\ t_0 = 2\,, \quad t_1 = t_2 &= 1\,. \end{aligned}$$

Es folgt:

$$\hat{a}(\xi) = \begin{vmatrix} 0 & \xi_1 & \xi_2 \\ \xi_1 & -1 & 0 \\ \xi_2 & 0 & -1 \end{vmatrix} = \xi_1^2 + \xi_2^2$$

$$\begin{aligned} \tilde{a} = a &= \hat{a}\ , \\ r = R &= 2\ . \end{aligned}$$

In der Theorie von Agmon, Douglis und Nirenberg wird weiter vorausgesetzt:

Bedingung L**:** $\tilde{a}(\mathbf{x}, \xi)$ ist ein Polynom vom Grade 2m bezüglich ξ. Sei ξ , $\hat{\xi} \in \mathbb{R}^n$ linear unabhängig. Das Polynom $q(\tau) := \tilde{a}(\mathbf{x}, \xi + \tau\hat{\xi})$ hat genau n Nullstellen τ mit positivem Imaginärteil.

Bemerkung 1.1 Sei das System elliptisch im Sinne von Douglis und Nirenberg und $n > 2$. Dann ist die Bedingung L erfüllt. (Siehe Agmon, Douglis, Nirenberg [1964, S. 39], Miranda [1970, S. 244].)

Bemerkung 1.2 Sei $n = 2$. Es gebe eine Funktion $\gamma(\mathbf{x})$ und ein $\alpha > 0$ mit

$$Re\{\gamma(\mathbf{x})\hat{a}(\mathbf{x}, \xi)\} \geq \alpha|\xi|^{2r} \quad \text{für alle} \;\; \xi \in \mathbb{R}^n\ ,\ \mathbf{x} \in \mathbb{R}^n\ ,\ \xi \neq 0\ ,$$

wobei $2r$ die Ordnung des Differentialoperators $\mathcal{M}_{ij}$ ist. Dann ist die Bedingung L erfüllt (siehe Miranda [1970, S. 245]).

Literatur

Agmon, S., **Douglis**, A., **Nirenberg**, L.: Estimates near the boundary for solutions of elliptic partial differential equations satisfying general boundary conditions. II. Comm. Pure Appl. Math. 17, 35-92(1964).

Douglis, A., **Nirenberg**, L.: Interior estimates for elliptic systems of partial differential equations. Comm. Pure Appl. Math. 8, 503-538(1955).

Hörmander, L.: Linear Partial Differential Operators. Berlin: Springer, 1963.

Miranda, C.: Partial Differential Equations of Elliptic Type. New York: Springer, 1970.

1.2 Randwertaufgaben

1.2.1 Einführung

Eine partielle Differentialgleichung hat mehrere Lösungen. In den Anwendungen werden Randbedingungen vorgeschrieben, die der Anwendung entsprechen und eine möglichst eindeutige Lösung bestimmen sollen. Für elliptische Gleichungen und Systeme gilt folgende Faustregel für sachgemäßgestellte Randwertaufgaben:

> Auf jedem Punkt des Randes $\partial\Omega$ von Ω sollten m Randbedingungen vorgeschrieben werden, wenn die Ordnung der Gleichung (des Systems) $2m$ ist.

Das folgende Beispiel von Hadamard zeigt, daß - wenn diese Faustregel verletzt wird - das Randwertproblem evtl. schlechtgestellt sein kann:

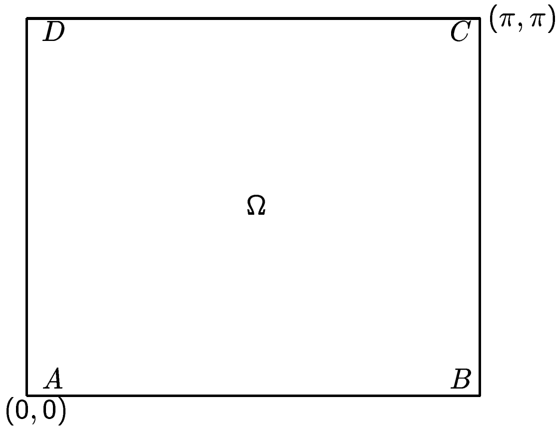

Abbildung 1.1: Das Beispiel von Hadamard

a) $\Delta u = 0\ ,\ (x,y) \in \Omega$

b) $u = 0$ auf AD

c) $u = 0$ auf BC

d) $u = 0$ auf AB

e) $u_n = -u_y = -\frac{1}{k^s}$ sin kx auf AB mit $k, s \in \mathbb{N}$.

Eine Lösung ist:

$$u(x,y) = \frac{1}{k^{s+1}} \cdot \sin\ kx \cdot \frac{e^{+ky} - e^{-ky}}{2}\ .$$

Für k großkann u beliebig großwerden, obwohl alle Randwerte gleichmäßig beschränkt sind.

1.2.2 Der Rand $\partial\Omega$

Der Rand $\partial\Omega$ mußeinige Glattheitsbedingungen erfüllen.

Typische Bedingungen sind:

1. Der Rand $\partial\Omega$ ist $C^{k,\alpha}$ mit $\alpha \in [0,1]$ und $k \in \mathbb{N}_0$, d.h. für jedes $\mathbf{x}_0 \in \partial\Omega$ gibt es eine Kugel $B = B(\mathbf{x}_0)$ und ein Homöomorphismus $\Psi : B \to D \subset \mathbb{R}^n$ mit

   (a) $\Psi(B \cap \Omega) \subset \mathbb{R}^n_+$

   (b) $\Psi(B \cap \partial\Omega) \subset \partial\mathbb{R}^n_+$

   (c)
   $$\begin{aligned} \Psi &\in C^{k,\alpha}(B)\,, \\ \Psi^{-1} &\in C^{k,\alpha}(D)\,, \end{aligned}$$

   wo
   $$\mathbb{R}^n_+ = \{\mathbf{x} \in \mathbb{R}^n : x_n > 0\}\,.$$

2. $\partial\Omega$ erfüllt eine gleichgradige innere Kegelbedingung, d.h. es gibt einen Kegel $K \subset \mathbb{R}^n$ mit der Spitze $\mathbf{0}$ und ein $r > 0$, so daß für jedes $x_0 \in \partial\Omega$ eine Kongruenz Ψ existiert mit
   $$x_0 + \Psi(K_r) \subset \Omega$$
   wo
   $$K_r := \{\mathbf{x} \in K : \|\mathbf{x}\| < r\}\,.$$
   (Siehe Abbildung 1.2.)

3. Ω sei ein Polygongebiet.

1.2.3 Typische Randbedingungen

1. <u>Eine Gleichung zweiter Ordnung</u>

   Sei
   $$Lu = \sum_{|\alpha| \le 2} a_\alpha(\mathbf{x})\partial^\alpha u$$
   eine elliptische Gleichung zweiter Ordnung. Seien $f, f_1, \ldots,$ vorgeschriebene Funktionen. Sei $\Omega \subset \mathbb{R}^n$. Es gibt drei kanonische Randwertaufgaben.

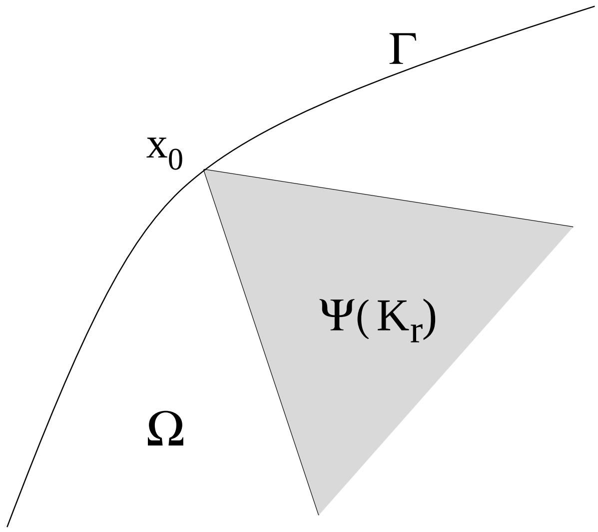

Abbildung 1.2: Die Kegelbedingung

(a) 1. Randwertaufgabe

$$\begin{aligned} Lu &= f\,, \quad \mathbf{x} \in \Omega \\ u &= f_1\,, \quad \mathbf{x} \in \partial\Omega \end{aligned}$$

Dieses Problem heißt auch *Dirichlet-Problem*.

(b) 2. Randwertaufgabe

$$\begin{aligned} Lu &= f\,, \quad \mathbf{x} \in \Omega \\ Lu &= \sum_{i,k=1}^{n} a_{ik}\,\frac{\partial^2 u}{\partial x_i \partial x_k} + \sum_{i=1}^{n} b_i\,\frac{\partial u}{\partial x_i} + c\,, \\ Lu &= f\,, \quad \mathbf{x} \in \Omega \\ \sum_{i,k=1}^{n} a_{ik}\nu_k\,\frac{\partial u}{\partial x_i} + f_2 u &= f_3\,, \quad \mathbf{x} \in \partial\Omega\,, \end{aligned}$$

wo $\nu = (\nu_1, \ldots, \nu_n)$ der äußere einheitsnormale n-Vektor auf $\partial\Omega$ ist.

Ist $f_2 = 0$, heißt dieses Problem auch *Neumann-Problem*, sonst *Robin-Problem*.

(c) 3. Randwertaufgabe

$$\begin{aligned} Lu &= f\,, \quad \mathbf{x} \in \Omega \\ f_4\,\frac{\partial u}{\partial \nu} + f_5 u &= f_6\,, \quad \mathbf{x} \in \partial\Omega\,. \end{aligned}$$

Dieses Problem heißt auch *gemischte Randwertaufgaben*.

2. Eine Gleichung vierter Ordnung

   Die bekannteste Gleichung vierter Ordnung ist die biharmonische Gleichung:

$$Lu = \frac{\partial^4 u}{\partial x^4} + 2\,\frac{\partial^4 u}{\partial x^2 \partial y^2} + \frac{\partial^4 u}{\partial y^4} = 0 \, .$$

   Typische Randbedingungen sind:

   (a) Das (verallgemeinerte) Dirichlet-Problem

$$\begin{aligned} Lu &= 0 \, , \quad \mathbf{x} \in \Omega \\ u &= f_1 \, , \quad \mathbf{x} \in \partial\Omega \\ u_n &= f_2 \, , \quad \mathbf{x} \in \partial\Omega \end{aligned}$$

3. Gleichungssysteme

   Siehe z.B. Agmon, Douglis, Nirenberg [1964, S. 42].

4. Existenz und Glattheit der Lösungen der Randwertaufgaben

   Für eine sachgemäßgestellte Randwertaufgabe für elliptische Gleichungen gilt folgende Faustregel:

   Die Lösung u ist so glatt wie die Koeffizienten der Gleichung, die Koeffizienten der Randbedingungen und der Rand es erlauben.

   Weiter gilt:

   Jede Änderung der Koeffizienten der Gleichung oder der Randbedingungen führt zu einer Änderung der Lösung u in jedem Punkt $\mathbf{x} \in \Omega$.

1.3 Existenz und Eindeutigkeit der Lösungen von Randwertaufgaben

Die Theorie von Randwertaufgaben benutzt mehrere Methoden, die die numerischen Methoden entweder anregen oder beeinflussen:

a) Funktionentheorie

b) Variationsgleichungen

c) Integralgleichungen

d) Fixpunktmethoden

e) Das Maximumprinzip

Bekannte numerische Methoden sind:

a) Differenzenverfahren

b) Finite Elemente

c) Randelementmethoden

d) Spektralmethoden

e) Galerkinmethoden

die oft mit

a) Mehrgittermethoden

b) Gebietszerlegungsmethoden

c) Gebietstransformationen

d) lokale Verfeinerung

kombiniert werden. Wir werden mehrere dieser Methoden und Techniken kennenlernen.

Kapitel 2

Das Maximumprinzip

Für elliptische Gleichungen zweiter Ordnung gilt ein Maximumprinzip, das außerordentlich nützlich ist.

2.1 Die grundlegenden Sätze

Sei

$$Mu := \sum_{i,k=1}^{n} a_{ik} \, \frac{\partial^2 u}{\partial x_i \partial x_k} + \sum_{i=1}^{n} b_i \, \frac{\partial u}{\partial x_i} \; .$$

Wir setzen voraus:

V1. $\Omega \subset \mathbb{R}^n$, Ω ein beschränktes Gebiet.

V2. $a_{ik} = a_{ik}(\mathbf{x})$, $b_i = b_i(\mathbf{x})$ seien stetige Funktionen auf Ω.

V3. Es gebe $\alpha > 0$ mit

$$\sum_{i,k=1}^{n} a_{ik}(\mathbf{x})\xi_i\xi_k \geq \alpha \sum_{i=1}^{n} \xi_i^2 \; , \quad \xi \in \mathbb{R}^n \; .$$

V4. $u \in C^2(\Omega) \cap C(\bar{\Omega})$.

Wir benutzen hier die Beweismethode von Courant und Hilbert, S. 326 ff.

Satz 2.1 *:*
u habe ein Maximum an der Stelle $P \in \Omega$. Dann gilt:

$$Mu(P) \leq 0 \; .$$

Beweis: Wenn u ein Maximum an dem inneren Punkt P hat, dann sind alle ersten Ableitungen von u an der Stelle P gleich Null. Folglich:

$$
\begin{aligned}
Mu(P) &= \sum_{i,k=1}^{n} a_{ik}(P)\,\frac{\partial^2 u}{\partial x_i \partial x_k} \\
&= \text{Spur} \quad AB
\end{aligned}
$$

mit

$$
\begin{aligned}
A &= (a_{ik}) \in \text{Mat}\,(n \times n)\,, \\
B &= \left(\frac{\partial^2 u}{\partial x_i \partial x_k}\,(P)\right) \in \text{Mat}\,(n \times n)\,.
\end{aligned}
$$

Es gibt eine orthogonale Matrix S mit

$$SAS^T = \Lambda = \text{Diag}\,(\lambda_i)\,.$$

Dann gilt:

$$
\begin{aligned}
\text{Spur}\,(AB) &= \text{Spur}\,(SABS^T) \\
&= \text{Spur}\,(\Lambda SBS^T) \\
&= \sum_{i=1}^{n} \lambda_i \beta_{ii}
\end{aligned}
$$

mit

$$\beta_{ii} = (SBS^T)_{ii}\,.$$

Gebe es i mit $\beta_{ii} > 0$, würde die Bedingung verletzt, daß u in P ein Maximum hat. Um dies einzusehen, setzt man $\xi = \epsilon \mathrm{s}$ mit

$$\mathrm{s} := \{s_{ij} : 1 \leq j \leq n\} \in \mathbb{R}^n\,.$$

Es gilt (Forster, II, Corollar 2, S. 59):

$$
\begin{aligned}
u(P+\xi) &= u(P) + \langle(\text{grad}\, u)(P)\,,\ \ \xi\rangle + \frac{1}{2}\langle\xi, B\xi\rangle + 0(\|\xi\|^2)\,, \\
&= u(P) + \frac{1}{2}\epsilon^2\langle \mathrm{s}, B\mathrm{s}\rangle + 0(\epsilon^2)\,, \\
&= u(P) + \frac{1}{2}\epsilon^2\beta_{ii} + 0(\epsilon^2)\,,
\end{aligned}
$$

so daß P kein lokales Maximum von u ist.

Die Voraussetzung V3 hat zur Folge, daß A positiv definit ist, so daß die Eigenwerte λ_i von A positiv sind.

Es folgt

$$Mu(P) = \sum_{i=1}^{n} \lambda_i \beta_{ii} \leq 0 \ .$$

Lemma 2.1 *(Hopf):*
Sei $K \subset \mathbb{R}^n$ eine offene Kugel und $P_0 \in \partial K$. Es gelte:

$$\begin{aligned} Mu \geq 0 \quad &, \quad \mathbf{x} \in K \\ u(\mathbf{x}) < u(P_0) \quad &, \quad \mathbf{x} \in K \ . \end{aligned}$$

Dann gilt:

$$\left.\frac{du}{dn}\right|_{p_0} := \liminf \left.\frac{\Delta u}{\Delta n}\right|_{p_0} > 0 \ .$$

Beweis: (siehe Abbildung 2.1)

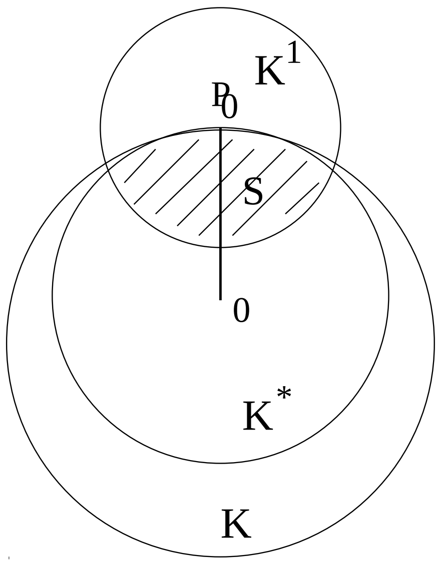

Abbildung 2.1: Hilfskonstruktion für den Beweis des Maximumprinzips

Sei K^* eine kleinere Kugel, $K^* \subset K$ mit $P_0 \in \partial K^*$. Dann gilt:

$$\max_{\bar{K}^*} u = u(P_0) \ .$$

Wähle den Ursprung O als Mittelpunkt von K^* und setze

$$r^2 := \sum_{i=1}^{n} x_i^2 \ .$$

Sei $r_0 = |P_0 - O|$. Sei K_1 eine Kugel mit Radius $r_1 < r_0$ und Mittelpunkt P_0. Setze

$$\begin{aligned} S &= K_1 \cap K^* \ , \\ \partial S &= R_1 \cup R_2 \ , \\ h(\mathbf{x}) &:= e^{-\beta r^2} - e^{-\beta r_0^2} \ , \quad \beta > 0 \ . \end{aligned}$$

Es gilt:

1. $h > 0 \ , \quad \mathbf{x} \in K^*$

2. $h = 0 \ , \quad \mathbf{x} \in \partial K^*$

3.

$$\begin{aligned} Mh &= e^{-\beta r^2} \cdot \left[4\beta^2 \sum_{i,k=1}^{n} a_{ik} x_i x_k - 2\beta \sum_{i=1}^{n} (a_{ii} + b_i x_i) \right] \\ &\geq e^{-\beta r^2} \left[4\beta^2 \alpha r^2 - 2\beta \sum_{i=1}^{n} a_{ii} - 2\beta r \sum_{i=1}^{n} b_i \right] \end{aligned}$$

Wähle $\beta > 0$ so, daß $M\,h > 0 \ , \quad \mathbf{x} \in \bar{S}$.

Sei $\epsilon > 0$ und $v := u + \epsilon h$.

Auf R_1 ist $u < u(P_0)$ und h beschränkt. Folglich gilt für genügend kleines ϵ:

$$v(\mathbf{x}) < u(P_0) \ , \quad \mathbf{x} \in R_1 \ .$$

Jetzt betrachten wir die Funktion v auf $\bar{S}$. Es gilt:

$$Mv = Mu + \epsilon Mh > 0 \ , \quad \mathbf{x} \in S \ .$$

Es folgt aus Satz 3.1

$$\max_{\bar{S}} \ v = \max_{\partial S} \ v \ .$$

Aber

$$\begin{aligned} v(\mathbf{x}) < u(P_0) &= v(P_0)\ , \quad \mathbf{x} \in R_1 \\ v(\mathbf{x}) &= u(\mathbf{x}) + \epsilon h(\mathbf{x}) \\ &= u(\mathbf{x}) \\ &< u(P_0)\ , \quad \text{für} \quad \mathbf{x} \in R_2 \backslash P_0\ . \end{aligned}$$

Folglich gilt:

$$\max_{\bar{S}}\ v = \max_{\partial S}\ v = v(P_0) = u(P_0)\ .$$

Es folgt:

$$\frac{dv}{dn}\,(P_0) = \lim_{\Delta n \to 0} \frac{v(P_0) - v(P)}{\Delta n} \geq 0\ .$$

Da

$$\frac{dh}{dn}\,(P_0) = -2\beta r_0 e^{-\beta r_0^2} < 0\ ,$$

folgt

$$\frac{du}{dn}\,(P_0) = \frac{dv}{dn}\,(P_0) - \epsilon\,\frac{dh}{dn}\,(P_0) > 0\ .$$

Satz 2.2 *(Starkes Maximumprinzip):*
Sei

$$Mu \geq 0\ , \quad \mathbf{x} \in \Omega\ .$$

Es gebe $P \in \Omega$ mit

$$\max_{\bar{\Omega}}\ u = u(P)\ .$$

Dann ist $u(\mathbf{x}) \equiv u(P)$, d.h. u ist konstant.

Beweis: Sei $Mu \geq 0$ in Ω. Falls u nicht konstant ist und es einen Punkt $P \in \Omega$ gibt mit

$$M := \max_{\bar{\Omega}}\ u = u(P)\ ,$$

dann gibt es eine offene Kugel K und einen Punkt P_0 mit

1. $\bar{K} \subset \Omega$
2. $u(\mathbf{x}) < u(P)\ , \quad \mathbf{x} \in K$
3. $P_0 \in \partial K$ mit $u(P_0) = u(P)$

Um dies einzusehen, sind folgende Überlegungen nötig:

1. Je zwei Punkte in einem zusammenhängenden Gebiet $\Omega \subset \mathbb{R}^n$ können mit einem Polygonzug $\Gamma \subset \Omega$ verbunden werden (Hocking und Young, Satz 3-5, S. 108).

2. Sei $P_1 \in \Omega$, $u(P_1) < M$. Sei $\Gamma \subset \Omega$ ein Polygonzug, der P und P_1 verbindet.

3. Es gilt:

$$d(\Gamma, \partial\Omega) = \rho > 0 \; .$$

4. Es gibt einen Punkt $P_2 \in \Gamma$ mit

$$\begin{aligned} d(P_2, \partial\Omega) &> d(P_2, \Omega_M) \; , \\ \Omega_M &:= \{\mathbf{x} \in \Omega : u(\mathbf{x}) = M\} \; . \end{aligned}$$

5. Es gibt eine Kugel $\tilde{K}$ mit Mittelpunkt P_2 und Punkt $P_0 \in \partial\tilde{K}$, die die gewünschten Eigenschaften besitzen.

Da u ein Maximum in P_0 hat, ist grad $u = 0$, was der Tatsache widerspricht, daß aus dem Hopfschen Lemma

$$\frac{du}{dn}\,(P_0) > 0 \; .$$

2.2 Existenz- und Eindeutigkeitsbeweise

Satz 2.3 *:*
Sei $\Omega \subset \mathbb{R}^n$ ein beschränktes Gebiet. Sei M der gleichmäßig elliptische Differentialoperator

$$M = \sum_{i,k=1}^{n} a_{ik}\, \frac{\partial^2 u}{\partial x_i \partial x_k} + \sum_{i=1}^{n} b_i\, \frac{\partial u}{\partial x_i} \; .$$

Dann hat das Dirichletsche Randwertproblem:

$$\begin{aligned} u &\in C^2(\Omega) \cap C(\Omega) \; , \\ Mu &= f \; , \quad \mathbf{x} \in \Omega \; , \\ u &= g \; , \quad \mathbf{x} \in \partial\Omega \end{aligned}$$

höchstens eine Lösung u.

Beweis: Anwendung des Maximumprinzips.
Es ist möglich, einen einfachen Existenzsatz für elliptische Gleichungen mit Hilfe des Maximumprinzips zu geben. Dies benutzt die Perronsche Methode von subharmonischen Funktionen. Wir verzichten auf den Beweis (siehe z.B. Gilbary und Trudinger [1977, S. 23]). Für die Laplace Gleichung erhält man z.B.:

Satz 2.4 *Sei $\Omega \subset \mathbb{R}^n$ ein beschränktes Gebiet, das die äußere Kugelbedingung erfüllt: Für jedes $\mathbf{x} \in \partial\Omega$ gibt es eine Kugel $K_R(\mathbf{y})$ mit:*

1. $K_R(\mathbf{y}) = \{\mathbf{x} : |\mathbf{x} - \mathbf{y}| < R\}$,
2. $\overline{K}_R(\mathbf{y}) \cap \overline{\Omega} = \{\mathbf{x}\}$.

Dann gibt es für jedes $g \in C(\partial\Omega)$ eine Funktion $u \in C^2(\Omega) \cap C(\overline{\Omega})$ mit

1. $u = g$ *auf* $\partial\Omega$
2. $\Delta u = 0$ *in* Ω.

Kürzlich sind diese Ideen erneut aufgenommen worden. Die sogenannten *viskosen Lösungen* ("viscous solutions") ermöglichen es, Existenz- und Eindeutigkeitssätze für viele nichtlineare partielle Differentialgleichungen zweiter Ordnung zu beweisen. Siehe Crandall, Ishii, Lions [1992].

Literatur:

Courant, R. und **Hilbert**, D.: Methods of Mathematical Physics II. Partial Differential Equations. New York: Interscience, 1962.

Crandall, M.G., **Ishii**, H. und **Lions**, P.-L.: Users guide to viscosity solutions of second order partial differential equations. Bull. Amer. Math. Soc. (New Series) 27 (1-67) 1992.

Gilbarg, D. und **Trudinger**, N.S.: Elliptic Partial Differential Equations of Second Order. Springer, 1977.

Hocking, J.G. und **Young**, G.S.: Topology. Reading MA, Addison-Wesley, 1961.

Protter, M.H. und **Weinberger**, H.F.: Maximum Principles in Differential Equations. Englewood Cliffs, NJ, Prentice-Hall, 1967.

Kapitel 3

Finite Differenzen

3.1 Einleitung

Das Differenzenverfahren oder die Methode der finiten Differenzen wird seit 1908 zur Lösung von Randwertaufgaben für partielle Differentialgleichungen benutzt. Sie verliert an Wichtigkeit im Vergleich zu neueren Methoden, wie Finiten Elementen, ist aber noch immer sehr nützlich. Viele Ideen werden von Differenzenverfahren auf die neueren Methoden übertragen.

Bahnbrechende Arbeiten wurden von Runge [1908], Richardson [1910] und Courant, Friedrichs und Lewy [1928] geleistet. Bekannte Standardwerke sind Forsythe und Wasow [1960], Collatz [1955]. Siehe auch Strikwerda [1989].

Sei $\Omega \subset \mathbb{R}^n$ ein Gebiet mit Rand $\Gamma = \partial\Omega$, seien A und B Differentialoperatoren. Eine allgemeine Randwertaufgabe für eine elliptische Differentialgleichung hat die Form: Bestimme u, so daß

$$\begin{aligned} Au &= f\ , \quad x \in \Omega\ , && (3.1)\\ Bu &= g\ , \quad x \in \Gamma\ . && (3.2) \end{aligned}$$

Es ist zuerst nötig, einige Ergänzungen zu machen:

a) Wichtig ist, die Voraussetzungen für die Lösung u anzugeben. Es wird meistens vorausgesetzt, daß u in einem vorgeschriebenen Banachraum V liegt. Ebenfalls gilt: $f \in F$.

b) Der Operator A mußals eine Abbildung zwischen den Banachräumen V und F definiert werden:

$$A : V \longrightarrow F\ .$$

c) Der Operator B und das Element g müssen auch sinnvoll definiert werden.

d) Die Randbedingungen werden oft als Teil der Definition von V betrachtet. Macht man dies, so mußnur die Gleichung

$$Au = f \tag{3.3}$$

gelöst werden.

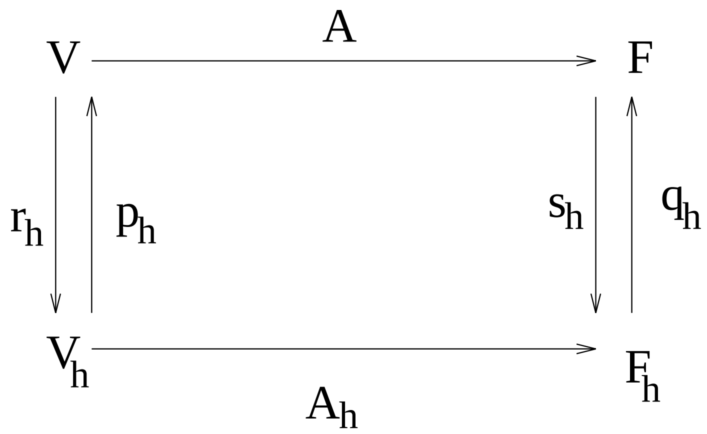

Abbildung 3.1: Interne Approximation

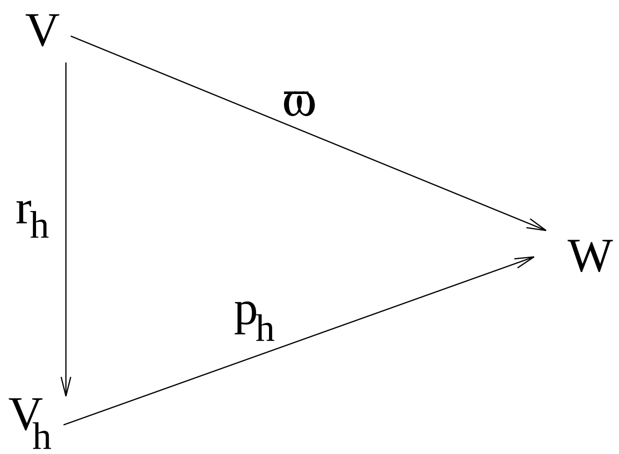

Abbildung 3.2: Externe Approximation

Die Gleichung $Au = f$ wird durch eine Gleichung

$$A_h u_h = f_h \tag{3.4}$$

approximiert, wobei

$$
\begin{aligned}
A_h \quad &: \quad V_h \longrightarrow F_h \ , \\
& f_h \in F_h
\end{aligned}
\tag{3.5}
$$

und V_h und F_h geeignete endlich dimensionale Banachräume sind.

Die Lösung der Gleichung

$$A_h u_h = f_h$$

erfordert üblicherweise die Lösung eines Systems linearer oder nichtlinearer algebraischer Gleichungen. Danach mußder Fehler zwischen u und u_h abgeschätzt werden. Es gibt drei Möglichkeiten:

1. u und $p_h u_h$ in V zu vergleichen (interne Approximation, Abbildung 3.1),
2. $\overline{\omega} u$ und $p_h u_h$ in einem anderen Raum W zu vergleichen (externe Approximation, Abbildung 3.2),
3. $r_h u$ und u_h in V_h zu vergleichen.

Die Beschreibungen "intern" und "extern" deuten an, daß die Approximation $p_h u_h$ innerhalb bzw. außerhalb des Lösungsraums V liegt.

Wichtig für jede Fehlerabschätzung sind zwei Begriffe:

Eine Approximation $A_h u_h$ von Au heißt *konsistent*, falls

$$||A_h r_h u - s_h A u||_{F_h} \longrightarrow 0 \quad \text{für} \quad h \longrightarrow 0 \ . \tag{3.6}$$

Eine Approximation A_h von A heißt *stabil*, falls eine Konstante $S > 0$ existiert mit

$$||v_h||_{V_h} \leq S ||A_h v_h||_{F_h} \ , \quad \text{für alle} \quad v_h \in V_h \ . \tag{3.7}$$

Die folgenden Fragen werden anschließend behandelt:

1. Wie werden die Abbildungen A_h konstruiert?
2. Existiert eine Lösung u_h der Gleichung $A_h u_h = f_h$?
3. Ist die Lösung u_h eindeutig?
4. Wie verhält sich der Fehler $u_h - r_h u$?

3.2 Ein Modellproblem: das Dirichlet-Problem für die Poisson- Gleichung

Wir betrachten das folgende Problem:

$$\Delta u = f\,, \quad (x,y) \in \Omega \tag{3.8}$$
$$u = g\,, \quad (x,y) \in \partial\Omega \tag{3.9}$$

wo Ω das Einheitsquadrat ist,

$$\Omega := \{(x,y) : 0 < x < 1\,,\ \ 0 < y < 1\}\,.$$

Dieses Problem heißt Dirichlet-Problem (wegen der Randbedingung (3.9)) für die Poisson-Gleichung (3.8), siehe Abbildung 3.3.

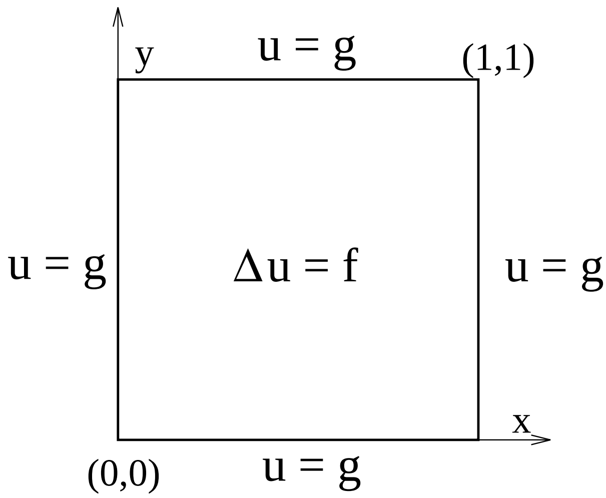

Abbildung 3.3: Das Dirichlet-Problem für die Poisson-Gleichung

Es wird eine Maschenweite $h = 1/M$ gewählt. Das Gebiet Ω wird mit einem quadratischen Gitter überzogen (siehe Abbildung 3.4).
Die Gitterpunkt sind die Punkte:

$$(x_i, y_j)\,,\ \ 0 \le i \le M\,,\ \ 0 \le j \le M\,, \tag{3.10}$$

wobei

$$\begin{aligned} x_i &= ih \\ y_j &= jh \end{aligned} \tag{3.11}$$

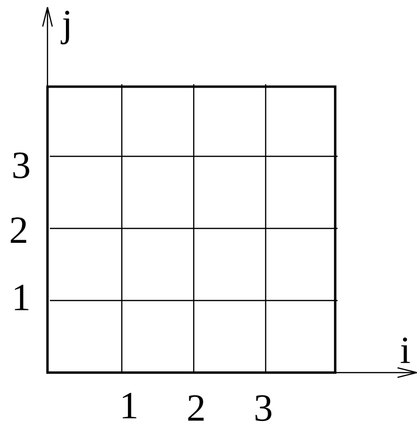

Abbildung 3.4: Das Gitter für $h = 1/4$

Da die Randbedingungen Dirichlet-Bedingungen sind, ist es nur nötig, die Werte von u_h im Innern des Quadrats zu bestimmen, und wir definieren:

$$\begin{aligned} \Omega_h &:= \{(x_i, y_j) : 1 \leq i\,, \quad j \leq M-1\} \\ \Gamma_h &:= \{(x_i, y_j) : i \in \{0, M\} \quad \text{oder} \quad j \in \{0, M\}\} \end{aligned} \tag{3.12}$$

Falls $u \in C^4(\overline{\Omega})$ und $(x, y) \in \Omega_h$ folgt:

$$\begin{aligned} \Delta_h u(x, y) &:= [u(x+h, y) + u(x-h, y) - 2u(x, y)] + \\ &+ [u(x, y+h) + u(x, y-h) - 2u(x, y)] \\ &= h^2 \Delta u + \frac{h^4}{12}[u_{xxxx}(\xi, y) + u_{yyyy}(x, \eta)]\,, \end{aligned}$$

mit

$$\xi \in (x-h, x+h)\,, \quad \eta \in (y-h, y+h)\,. \tag{3.13}$$

Somit erhält man als Approximation zu den Gleichungen (3.8), (3.9):

$$\Delta_h u_h(x_i, y_j) = f(x_i, y_j)\,, \quad (x_i, y_j) \in \Omega_h \tag{3.14}$$

$$u_h(x_i, y_j) = g(x_i, y_j)\,, \quad (x_i, y_j) \in \Gamma_h \tag{3.15}$$

Für dieses Modellproblem gilt:

$$\begin{aligned} V &= C^4(\Omega) \cap C(\overline{\Omega}) \\ F &= C^2(\Omega) \\ Au &= \Delta u \end{aligned} \tag{3.16}$$

$$
\begin{aligned}
(A_h u_h)(x_i, y_j) &:= \Delta_h u_h(x_i, y_j) \\
V_h &:= \prod_{k=1}^{N} \mathbb{R}^1 = \mathbb{R}^N \\
N &:= (M-1)^2 \\
F_h &:= \mathbb{R}^N \\
r_h &: C(\Omega) \longrightarrow F_h \\
(r_h u)(x_i, y_j) &:= u(x_i, y_j)\ .
\end{aligned}
\tag{3.17}
$$

(r_h heißt Injektionsoperator.)

Der Differenzenoperator Δ_h läßt sich wie in Abbildung 3.5 darstellen und heißt deshalb das Fünfpunkt-Stern oder Fünfpunkt-Molekül.

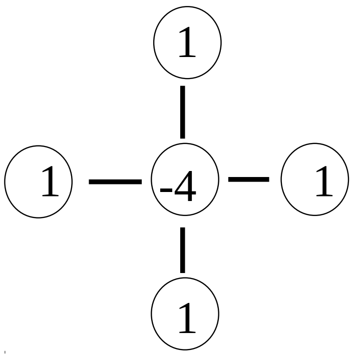

Abbildung 3.5: Das Fünfpunkt-Molekül

A_h ist die $(M-1)^2 \times (M-1)^2$-Blocktridiagonal-Matrix:

$$
A_h := \begin{pmatrix} J & I & & \\ I & \ddots & \ddots & \\ & \ddots & \ddots & I \\ & & I & J \end{pmatrix}
$$

$$
I = I_{M-1} = \begin{pmatrix} 1 & & \\ & \ddots & \\ & & 1 \end{pmatrix}
$$

$$
J = J_{M-1} = \begin{pmatrix} -4 & 1 & & \\ 1 & \ddots & \ddots & \\ & \ddots & \ddots & 1 \\ & & 1 & -4 \end{pmatrix}
$$

3.3 Differenzenoperatoren

Es gibt ein Kalkül für Differenzenoperatoren, das oft nützlich ist bei der Erstellung von Differenzengleichungen.

Sei $f \in C^\infty(-\infty, +\infty)$ und $h \in \mathbb{R}$. Wir führen folgende Operatoren ein:

Ef	$=$	$f(x+h)$	Verschiebungsoperator
$\Delta f(x)$	$=$	$f(x+h) - f(x)$	Vorwärtsdifferenzenoperator
$\nabla f(x)$	$=$	$f(x) - f(x-h)$	Rückwärtsdifferenzenoperator
$Df(x)$	$=$	$f'(x)$	Differentiationsoperator
$\delta f(x)$	$=$	$f\left(x + \frac{1}{2}h\right) - f\left(x - \frac{1}{2}h\right)$	Zentraler Differenzenoperator
$\mu f(x)$	$=$	$\frac{1}{2}\left[f\left(x + \frac{1}{2}h\right) + f\left(x - \frac{1}{2}h\right)\right]$	Mittelwertoperator

Die Operatoren sind linear, assoziativ und kommutativ. Sie bilden einen Ring. Es folgt, daß der binomische Satz gilt:

$$(P+Q)^n = \sum_{k=0}^{n} P^k Q^{n-k} \binom{n}{k} .$$

Aus der Taylor-Reihe

$$Ef(x) = f(x+h) = \sum_{k=0}^{\infty} \frac{h^k}{k!} D^k f(x)$$

folgt:

$$E = e^{hD} . \tag{3.18}$$

Also ist

$$e^{hD} = 1 + \Delta . \tag{3.19}$$

Logarithmiert man diese Gleichung, so erhält man:

$$\begin{aligned} hD &= \ell n(1+\Delta) \\ &= \Delta - \frac{\Delta^2}{2} + \frac{\Delta^3}{3} \cdots \end{aligned} \tag{3.20}$$

Man kann diese Gleichungen folgendermaßen beweisen:

Sei $f(x) = e^{ax}$. Dann gilt

$$\sum_{k=1}^{n} \frac{(-1)^k}{k} \Delta^k f(x) = \sum_{k=1}^{n} \frac{(-1)^k}{k} (e^{ah} - 1)^k f(x) \, .$$

Für $|e^{ah} - 1| < 1$ folgt:

$$\sum_{k=1}^{\infty} \frac{(-1)^k}{k} \Delta^k f(x) = \ell n(1 + (e^{ah} - 1)) f(x) = ahe^{ax} = hDf(x) \, .$$

Also gilt:

$$\ell n(1 + \Delta) f(x) = hDf(x) \qquad (*)$$

für $f(x) = e^{ax}$ und $|e^{ah} - 1| < 1$.

Nun setzt man

$$e^{ax} = \sum_{\ell=0}^{\infty} \frac{(ax)^\ell}{\ell!}$$

ein, entwickelt die beiden Seiten von $(*)$ nach Potenzen von a und vergleicht die Koeffizienten von $\frac{a^\ell}{\ell!}$. Es folgt:

$$hDx^\ell = \sum_{k=1}^{\ell} \frac{(-1)^k}{k} \, \Delta^k x^\ell \, ,$$

d.h. (3.20) gilt für alle Polynome.

Weitere Formeln sind:

$$\begin{aligned}
\delta &= 2 \sinh \left(\frac{1}{2} hD \right) \\
hD &= 2 \sinh^{-1} \left(\frac{\delta}{2} \right) \\
&= 2 \left[\frac{\delta}{2} - \frac{1}{2} \left(\frac{\delta}{2} \right)^3 \frac{1}{3} + \frac{1}{2} \cdot \frac{3}{4} \left(\frac{\delta}{2} \right)^5 \frac{1}{5} - \frac{1 \cdot 3 \cdot 5}{2 \cdot 4 \cdot 6} \left(\frac{\delta}{2} \right)^7 \frac{1}{7} \cdots \right] \\
&= \delta - \frac{\delta^3}{24} + \frac{3\delta^5}{640} - \frac{5\delta^7}{7168} \cdots \\
(hD)^2 &= \delta^2 - \frac{\delta^4}{12} + \frac{\delta^6}{90} - \frac{\delta^8}{560} \cdots \\
\mu \left(1 + \frac{1}{4} \delta^2 \right)^{-1/2} &= 1 \\
hD &= \mu\delta - \mu \frac{\delta^3}{6} + \mu \frac{\delta^5}{30} \cdots
\end{aligned}$$

Diese Formeln werden in Björck und Dahlquist [1979, §7.6] und Nörlund [1954] bewiesen. Man kann diese Formeln natürlich auch als Hilfsmittel betrachten, die es ermöglichen, Differenzenapproximationen schnell herzuleiten, wobei die Richtigkeit nachher mit Hilfe von Taylor-Reihenentwicklungen bestätigt werden muß.

Beispiel 3.1 a)

$$\begin{aligned} u_x = Du &= \frac{1}{h}\ell n(1+\Delta)u \\ &= \frac{1}{h}\left[\Delta u - \frac{\Delta^2 u}{2}\cdots\right] \\ \Rightarrow u_x(x) &\doteq \frac{1}{h}\Delta u(x) \\ &= \frac{1}{h}[u(x+h)-u(x)]\ . \end{aligned}$$

b) $u_x(x) \doteq \frac{1}{h}\nabla u(x) = \frac{1}{h}[u(x)-u(x-h)]$.

c)

$$\begin{aligned} u_x(x) &= \frac{2}{h}\sinh^{-1}(\delta/2)u(x) \\ &= \frac{2}{h}\sinh^{-1}(\delta/2)\left(1+\frac{1}{4}\delta^2\right)^{-1/2}\mu u(x) \\ &\doteq \frac{2}{h}\cdot\frac{\delta}{2}\cdot\mu\, u(x) \\ &= \frac{u(x+h)-u(x-h)}{2h}\ . \end{aligned}$$

d)

$$\begin{aligned} u_{xx} &= \frac{4}{h^2}\left(\sinh^{-1}\delta/2\right)^2 u \\ &= \frac{4}{h^2}\left(\frac{\delta}{2}-\frac{\delta^3}{48}\cdots\right)^2 u \\ &= \frac{4}{h^2}\frac{\delta^2}{4}\left(1-\frac{\delta^2}{24}\cdots\right)^2 u \\ &= \frac{\delta^2}{h^2}\left(1-\frac{\delta^2}{12}\cdots\right)u\ . \end{aligned}$$

Das bedeutet, daß die Differentialgleichung

$$u''(x) = f(x)$$

durch

$$\frac{1}{h^2}\left(\delta^2 - \frac{1}{12}\delta^4\right) u_i = f_i$$

oder

$$\frac{1}{h^2}\left(-\frac{1}{12}u_{i-2} + \frac{4}{3}u_{i-1} - \frac{5}{2}u_i + \frac{4}{3}u_{i+1} - \frac{1}{12}u_{i+2}\right) = f_i$$

approximiert werden kann. Diese Approximation ist zwar genauer als die Standardapproximation, hat aber mehrere Nachteile: Verlust der Positivität; 5-Diagonalmatrizen; usw.

Im allgemeinen besteht das Differenzenverfahren daraus, daß alle partiellen Ableitungen in einer partiellen Differentialgleichung durch Differenzen ersetzt werden. Es gibt oft mehrere Möglichkeiten, unter denen eine Auswahl getroffen werden muß.

3.4 Finite Differenzen: Fortsetzung

Das Differenzenverfahren für das Problem

$$\begin{aligned} Au &= f\,, \text{ in } \Omega \\ Bu &= g\,, \text{ auf } \partial\Omega \end{aligned} \tag{3.21}$$

besteht aus den folgenden Schritten:

a) $\mathbb{R}^n$ wird mit einem regelmäßigen Gitter G_h überzogen. $\Omega_h := G_h \cap \Omega$ heißt die Menge der *inneren Gitterpunkte*. Am Rand werden weitere Punkte ausgewählt, die meistens die Schnittpunkte von Gitterlinien und $\partial\Omega$ sind. Die Menge dieser Punkte sind die *Randgitterpunkte*. Sie wird mit Γ_h bezeichnet. Siehe Abbildung 3.6.

   Der Raum von Gitterfunktionen

   $$V_h := C^0(\Omega_h \cup \Gamma_h, \mathbb{R})$$

   ist die Menge aller reellen Funktionen, die auf $\Omega_h \cup \Gamma_h$ definiert sind. Die Bezeichnung C^0 deutet an, daß die Funktionen stetig sind, was in diesem Fall keine Einschränkung ist, da $\Omega_h \cup \Gamma_h$ eine diskrete Menge ist.

b) Sei $N = N_h$ die Dimension von V_h,

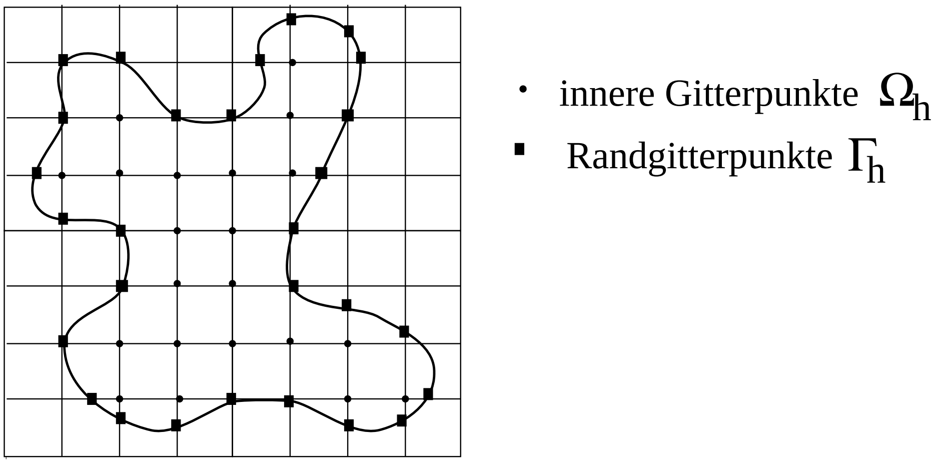

Abbildung 3.6: Die Gitterpunktmengen

$$N_h = |\Omega_h| + |\Gamma_h| \ .$$

Ein System von N_h Gleichungen,

$$\begin{aligned} A_h(v_h) &= f_h \\ A_h &: V_h \longrightarrow F_h \\ f_h &:= sf \in F_h \end{aligned} \tag{3.22}$$

wird so konstruiert, daß die Restriktion von u auf Ω_h

$$u_h := r_h u \tag{3.23}$$

das System (3.22) mit geringem Fehler erfüllt:

$$A_h(u_h) - f_h = 0(h^m) \ . \tag{3.24}$$

Das Gleichungssystem (3.22) wird so konstruiert, daß jedem Gitterpunkt $P \in \Omega_h \cup \Gamma_h$ eine Gleichung $A_h^P(v_h) = f_h(P)$ zugeordnet wird:

1. Für $P \in \Omega_h$ entsteht diese Gleichung dadurch, daß in der Gleichung

   $$(Au)(P) = f(P)$$

   die partiellen Ableitungen durch Differenzen ersetzt werden.
2. Für $P \in \Gamma_h$ werden auch partielle Ableitungen durch Differenzen approximiert, wobei allerdings sehr viele Möglichkeiten berücksichtigt werden müssen.

3.4.1 Beispiel 1: Das Neumann Problem

$$\begin{aligned} \Delta u &= 0\,, \text{ in } \Omega \\ \frac{\partial u}{\partial n} &= g\,, \text{ auf } \Gamma \end{aligned} \tag{3.25}$$

Sei $u \in C^2(\overline{G})$ und Γ glatt. Es folgt aus dem Gaußschen Integralsatz, daß

$$0 = \int_\Omega div \;\; grad \;\; u \, dx = \int_\Gamma \frac{\partial u}{\partial n} ds = \int_\Gamma g \, ds \,.$$

Das Problem (3.25) ist deshalb nur sinnvoll, wenn g die folgende Bedingung erfüllt:

$$\int_\Gamma g \, ds = 0 \,. \tag{3.26}$$

Wird diese Bedingung erfüllt und erfüllt u die Gleichung (3.25), dann ist
$v := u + \text{const.}$ auch eine Lösung von (3.25).

Wir betrachten jetzt das Neumann Problem (3.25) für das Rechteck $\Omega = (0,1)^2$ mit $h = 1/2$ (Abbildung 3.7).

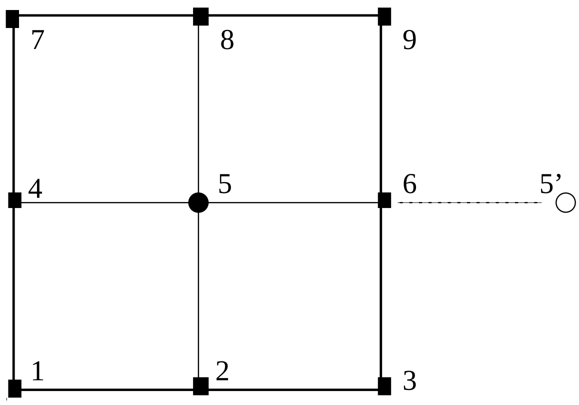

Abbildung 3.7: Das Neumann Problem auf einem Rechteck

Für den Randgitterpunkt P_6 gelten folgende Überlegungen: Wäre P_6 ein innerer Punkt mit Nachbarn $P_3, P_5, P_5^{'}$ und P_9 (siehe Abbildung 3.7), dann würde gelten:

$$u_5 + u_{5'} + u_9 + u_3 - 4u_6 \doteq \Delta u(P_6) = 0 \,.$$

Die Randbedingung $\frac{\partial u}{\partial n} = g_6$ kann durch die Gleichung

$$\frac{u_5^{'} - u_5}{2h} = g_6$$

approximiert werden. Elimination von u_5 führt zu der Gleichung

$$2u_5 + u_9 + u_3 - 4u_6 = -2hg_6 \ ,$$

die selbstverständlich auch durch eine Taylorreihenentwicklung überprüft werden kann.

Entsprechende Überlegungen bei den übrigen Randgitterpunkten liefert folgendes Gleichungssystem:

$$A_h v_h = f_h$$

mit

$$v_h = (v_1, \ldots, v_9)^T$$

$$A_h = \left(\begin{array}{ccc|ccc|ccc}
-4 & 2 & 0 & 2 & 0 & 0 & 0 & 0 & 0 \\
1 & -4 & 1 & 0 & 2 & 0 & 0 & 0 & 0 \\
0 & 2 & -4 & 0 & 0 & 2 & 0 & 0 & 0 \\
\hline
1 & 0 & 0 & -4 & 2 & 0 & 1 & 0 & 0 \\
0 & 1 & 0 & 1 & -4 & 1 & 0 & 1 & 0 \\
0 & 0 & 1 & 0 & 2 & -4 & 0 & 0 & 1 \\
\hline
0 & 0 & 0 & 2 & 0 & 0 & -4 & 2 & 0 \\
0 & 0 & 0 & 0 & 2 & 0 & 1 & -4 & 1 \\
0 & 0 & 0 & 0 & 0 & 2 & 0 & 2 & -4
\end{array}\right)$$

$$f_h = -2h \left(g_1^{(x)} + g_1^{(y)}, g_2, g_3^{(x)} + g_3^{(y)}, g_4, 0, g_6, g_7^{(x)} + g_7^{(y)}, g_8, g_9^{(x)} + g_9^{(y)} \right)^T \ ,$$

wobei die Terme $g_1^{(x)} + g_1^{(y)}$ dadurch entstehen, daß an Eckpunkten zwei äußere Normale existieren.

Die Matrix A_h ist, wie zu erwarten, singulär: die Zeilensummen sind null.

Dieses Beispiel zeigt, wie leicht es ist, Differenzengleichungen für Probleme zu erstellen, für die Existenzsätze nicht sofort vorhanden sind.

3.4.2 Beispiel 2: Nichtquadratische Gitter

Es ist oft zweckmäßig, nichtquadratische Gitter zu benutzen:

1. Sei $\Omega = [0,a] \times [0,b]$, wobei a/b eine irrationale Zahl sei. Dann gibt es keine Maschenweite h mit

$$a = mh \ , \quad b = nh \ , \quad m, n \in \mathbb{N} \ .$$

2. Sei $\varphi : [0,1] \longrightarrow [0,1]$, $\varphi(0) = 1$, $\varphi(1) = 0$, φ monoton fallend.

Das Gebiet Ω

$$\Omega = \{(x,y) : 0 \leq x \leq 1 \ , \quad 0 \leq y \leq \varphi(x)\}$$

wird in Abbildung 3.8 gezeigt.

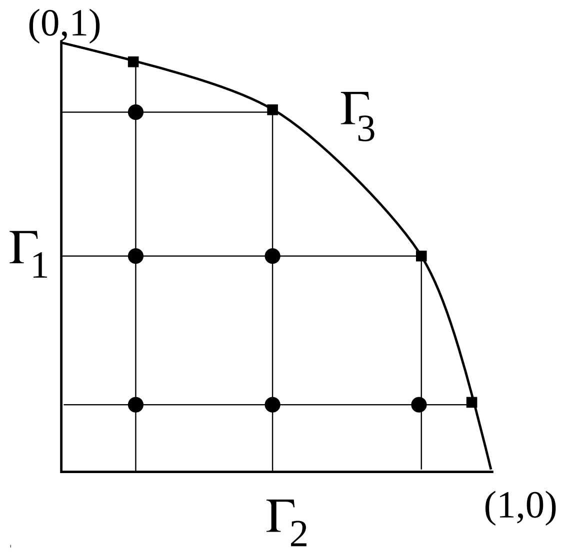

Abbildung 3.8: Ein nichtquadratisches Gitter

Es werden zuerst die Randgitterpunkte auf Γ_3 gewählt (siehe Abbildung 3.8), dann ergibt sich ein geeignetes Gitter.

In einem allgemeinen Gitterpunkt (x_i, y_j) erhält man z.B. mit der Schreibweise von Abbildung 3.9:

$$\begin{aligned}
\frac{1}{h^2}\left[\frac{u_{i+1,j} - u_{i,j}}{s_1} - \frac{u_{i,j} - u_{i-1,j}}{s_3}\right]\frac{2}{s_1 + s_3} &= u_{xx}(x_i, y_j) - R_{ij} \ , \\
R_{ij} = \frac{(s_3 - s_1)h}{3} u_{xxx}(x_i, y_j) &- \frac{s_3^2 - s_3 s_1 + s_1^2}{12} h^2 u_{xxxx}(\xi, y_j) \ , \\
& \quad x_j - s_3 h < \xi < x_j + s_1 h \ .
\end{aligned}$$

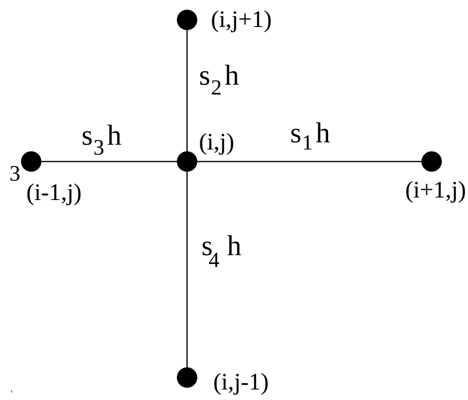

Abbildung 3.9: Ein allgemeiner Gitterpunkt

Falls

$$
\begin{aligned}
|s_k| \leq 1 \quad &, \quad 1 \leq k \leq 4\,, \\
M_3 \quad &:= \quad \sup_{\Omega} |u_{xxx}| \\
M_4 \quad &:= \quad \sup_{\Omega} |u_{xxxx}| \quad \text{gilt} \\
|R_{ij}| \leq \frac{1}{3} h M_3 \quad &+ \quad \frac{1}{12} h^2 M_4
\end{aligned}
$$

Siehe z.B. Forsythe und Wasow [1960, S. 188].

3.4.3 Beispiel 3: Neumann Bedingungen auf gekrümmten Rändern

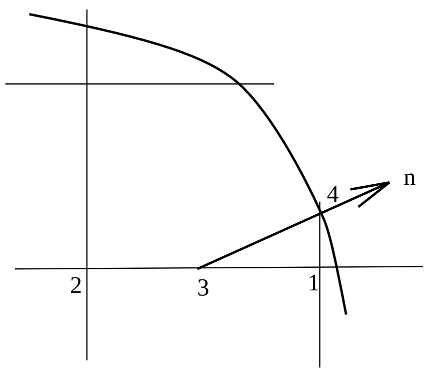

Abbildung 3.10: Approximation von Neumann Bedingungen

Die Normalenableitung am Punkt P_4 in Abbildung 3.10 kann folgendermaßen approximiert werden:

$$\frac{\partial u}{\partial n}(P_4) \doteq \frac{u_4 - \tilde{u}_3}{|P_4 - P_3|} ,$$

wobei

$$\tilde{u}_3 = \frac{u_1 \cdot |P_3 - P_2| + u_2 |P_1 - P_3|}{|P_1 - P_2|} .$$

3.4.4 Beispiel 4:

Die Minimaloberflächengleichung

$$(1 + u_y^2)u_{xx} - 2u_x u_y u_{xy} + (1 + u_y^2)u_{xx} = 0$$

kann folgendermaßen auf einem quadratischen Gitter (Abbildung 3.11) approximiert werden:

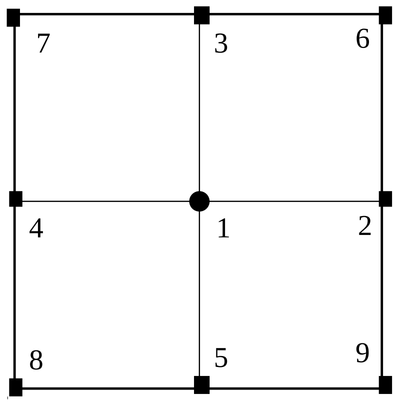

Abbildung 3.11: Nachbarpunkte für die Minimaloberflächengleichung

$$\begin{aligned}
u_x &\doteq (u_2 - u_4)/2h \\
u_y &\doteq (u_3 - u_5)/2h \\
u_{xx} &= (u_2 + u_4 - 2u_1)/h^2 \\
u_{yy} &= (u_3 + u_5 - 2u_1)/h^2 \\
u_{xy} &= (u_6 + u_8 - u_7 - u_9)/4h^2 .
\end{aligned}$$

3.5 Das Maximumprinzip für Differenzengleichungen

Das Maximumprinzip für elliptische partielle Differentialgleichungen läßt sich auf Differenzengleichungen übertragen.

Sei G_h ein Gitter und $S \subset G_h$. Ein linearer Differenzenoperator ℓ, z.B. Δ_h, kann folgendermaßen geschrieben werden:

$$(\ell u)(P) = \sum_{Q \in S} A(P,Q)u(Q) \ . \tag{3.27}$$

Man sagt, der Differenzenoperator ℓ sei nichtnegativ an der Stelle $P \in S$, falls

$$(\ell u)(P) = \sum_{Q \in S} A(P,Q)u(Q) \ , \tag{3.28}$$

$$A(P,P) < 0 \ , \quad A(P,Q) \geq 0 \quad \text{für} \ \ P \neq Q \ , \tag{3.29}$$

und

$$\sum_{Q \neq P} |A(P,Q)| \leq |A(P,P)| \ . \tag{3.30}$$

Man setzt

$$\begin{aligned} (\ell u)(P) &= \sum_{Q \in S} B(P,Q)u(Q) + f(P)u(P) \\ &= (mu)(P) + f(P)u(P) \end{aligned} \tag{3.31}$$

mit

$$\sum_{Q \in S} B(P,Q) = 0 \ . \tag{3.32}$$

In diesem Absatz wird stets angenommen, daß m nichtnegativ ist, d.h.

$$B(P,P) < 0 \ , \tag{3.33}$$

$$B(P,Q) \geq 0 \ , \quad Q \neq P \ , \tag{3.34}$$

$$\sum_{Q \neq P} B(P,Q) = -B(P,P) \ . \tag{3.35}$$

Die Menge der Nachbarn von $P \in G_h$ bzgl. ℓ wird durch

$$\mathcal{N}(P) := \{Q \in G_h : A(P,Q) \neq 0 \ , \ Q \neq P\} \tag{3.36}$$

bezeichnet.

Ist R eine Teilmenge von S, wird ∂R, der Rand von R bzgl. ℓ, erklärt durch

$$\begin{aligned} \partial R &:= \{Q \in S - R : A(P,Q) \neq 0 \text{ für mindestens einen Punkt } P \in R\} \\ \overline{R} &:= R \cup \partial R\,. \end{aligned} \tag{3.37}$$

Eine Menge $R \subset G_h$ heißt ***zusammenhängend bzgl.*** ℓ, falls es für jedes Paar $P, Q \in R$ eine endliche Folge $P_1, \ldots, P_r$ gibt mit:

$$\begin{aligned} P_1 &= P \\ P_r &= Q \\ A(P_j, P_{j+1}) &\neq 0\,, \quad 1 \leq j \leq r-1\,. \end{aligned} \tag{3.38}$$

Sei $H = H(S)$ die Menge aller Gitterfunktionen auf S.

Es gilt

$$\begin{aligned} (mv)(P) &= \sum_{Q \in S} B(P,Q) v(Q) \\ &= \sum_{Q \neq P} B(P,Q)[v(Q) - v(P)] + v(P) \sum_{Q \in S} B(P,Q) \\ &= \sum_{Q \neq P} B(P,Q)[v(Q) - v(P)] \end{aligned} \tag{3.39}$$

Man sagt, daß $v \in H$ ein lokales Maximum/Minimum bzgl. ℓ an der Stelle P hat, falls

$$\begin{gathered} v(Q) \leq v(P) \quad \text{für} \quad Q \in \mathcal{N}(P) \\ \text{bzw.} \\ v(Q) \geq v(P) \quad \text{für} \quad Q \in \mathcal{N}(P)\,. \end{gathered} \tag{3.40}$$

Lemma 3.1 *$v \in H$ habe ein lokales Maximum an der Stelle P. Dann gilt* $(mv)(P) \leq 0$.

Beweis:

$$\begin{aligned} (mv)(P) &= \sum_{Q \neq P} B(P,Q)[v(Q) - v(P)] \\ &\leq 0 \end{aligned}$$

wegen (3.40) und (3.39).

Satz 3.1 (Das Maximumprinzip) *Sei R eine zusammenhängende Teilmenge von S.*

1. *Falls $mv \geq 0$ in R und v ein Maximum auf $\overline{R}$ in einem Punkt $P \in R$ annimmt, dann ist v konstant auf $\overline{R}$.*

2. *Sei $f \leq 0$ und $\ell v \geq 0$ in R, v nehme sein Maximum auf $\overline{R}$ in einem inneren Punkt an, und dieses Maximum sei positiv. Dann ist v konstant auf $\overline{R}$.*

3. *Sei $f \leq 0$ und $\ell v \geq 0$ in R und $\partial R \neq \phi$. Dann gilt*

$$v \leq \max\{0, \max_{\partial R} v\}$$

4. *Sei $mv \geq 0$ in R und das Maximum von v auf $\overline{R}$ in einem Punkt $P_0 \in \partial R$ erreicht. Dann ist entweder $v \equiv$ konstant oder $v(P_0) - v(P) > 0$ für $P \in R$.*

5. *Sei $f \leq 0$ und $\ell v \geq 0$ in R und sei das Maximum von v auf $\overline{R}$ in einem Punkt $P_0 \in \partial R$ erreicht. Ist das Maximum positiv, dann ist entweder $v \equiv$ konstant oder $v(P_0) - v(P) > 0$ für $P \in R$.*

Beweis von 1:

Sei $P \in R$ mit

$$v(P) = \mu = \max_{\overline{R}} v \ .$$

Da v an der Stelle das Maximum erreicht, folgt aus Lemma 3.1, daß $(mv)(P) \leq 0$. Es wurde angenommen, daß $mv(P) \geq 0$. Es folgt $mv(P) = 0$, d.h. $(mv)(P) = \sum_{Q \neq P} B(P,Q)[v(Q) - v(P)] = 0$. Folglich gilt

$$v(Q) = v(P) = \mu \ , \quad Q \in \mathcal{N}(P) \ .$$

Aber $\overline{R}$ ist ℓ-zusammenhängend, so daß $v(Q) = \mu$ für alle $Q \in \overline{R}$.

Beweis von 2:

Sei $P \in R$ und $v(P) = \mu = \max_{\overline{R}} v > 0$. Es folgt - wie in dem Beweis von 1 - daß $mv(P) \leq 0$. Es wurde vorausgesetzt, daß $\ell v(P) \geq 0$. Es ergibt sich

$$mv(P) = \ell v(P) - f(P)v(P) \geq 0 \ ,$$

so daß $mv(P) = 0$. Wie bei dem Beweis von 1 folgt jetzt, daß $v(Q) = \mu$ für $Q \in \overline{R}$.

Beweis von 3:

3 ist nur eine andere Formulierung von 2.

Beweis von 4 und 5:

Sofort einzusehen.

Satz 3.2 *Sei R eine ℓ-zusammenhängende Teilmenge von S und $\partial R \neq \phi$. Sei ℓ nichtnegativ auf R. Die Funktionen γ und g haben Definitionsbereiche R bzw. ∂R. Dann hat das Problem*

$$\begin{aligned} \ell v &= g\,, \quad \textit{in } R \\ v &= \gamma\,, \quad \textit{auf } \partial R \end{aligned} \tag{3.41}$$

eine eindeutige Lösung $v \in H(\overline{R})$.

Beweis:

Nach Voraussetzung gilt

$$\ell v(P) = mv(P) + f(P)v(P)$$

mit $f \leq 0$.

Das Problem (3.41) ist auch ein Problem in linearer Algebra: es gibt $M = |\overline{R}|$ Gleichungen für die M Unbekannten $v(P),\ P \in \overline{R}$. Es folgt, daß v existiert, falls v eindeutig ist, d.h. falls $V \equiv 0$ die einzige Lösung des folgenden Problems ist:

$$\begin{aligned} \ell V &= 0\,, \quad \text{in } R \\ V &= 0\,, \quad \text{auf } \partial R \end{aligned} \qquad *$$

Aber mit Hilfe des Satzes 3.1 sehen wir, daß

$$\begin{aligned} \max_{\overline{R}} V \leq \max\{0,0\} &= 0\,, \\ \max_{\overline{R}} (-V) \leq \max\{0,0\} &= 0\,, \end{aligned}$$

so daß $V \equiv 0$ wie benötigt.

Satz 3.3 *Sei R eine ℓ-zusammenhängende Teilmenge S mit $\partial R \neq \phi$.Sei ℓ nichtnegativ auf R. Sei $\phi \in H(\overline{R})$ mit*

$$\begin{aligned} \ell\phi &\geq 1\,, \quad \textit{in } R \\ \phi &\geq 0\,, \quad \textit{auf } \overline{R} \end{aligned} \tag{3.42}$$

Sei

$$M = \max_{\partial R} \phi\,. \tag{3.43}$$

Dann gilt für alle $v \in H$:

$$\max_{\overline{R}} |v(P)| \leq \max_{\partial R} |v(P)| + M \max_{R} |\ell v| \tag{3.44}$$

Beweis:

Sei

$$w_{\pm} = v \pm \phi \max_{R} |\ell v|\,. \tag{3.45}$$

Man rechnet nach, daß

$$\ell w_{-} \leq 0 \leq \ell w_{+} \text{ in } R\,.$$

Folglich (siehe Satz 3.1):

$$\begin{aligned} w_{+} &\leq \max\{0, \max_{\partial R} w_{+}\} \\ w_{-} &\geq \min\{0, \min_{\partial R} w_{-}\}\,, \text{ in } \overline{R} \end{aligned} \tag{3.46}$$

Aus (3.45) und (3.46) mit Hilfe von (3.42) und (3.43) schließt man:

$$\begin{aligned} v \leq w_{+} &\leq \max_{\partial R} |w_{+}| \\ &\leq \max_{\partial R} |v| + M \max_{R} |\ell v|\,. \end{aligned}$$

Genauso folgt:

$$v \geq w_{-} \geq -\max_{\partial R} |w_{-}| \geq -[\max_{\partial R} |v| + M \max_{R} |\ell v|]\,.$$

Damit ist der Satz bewiesen.

3.6 Konvergenzbeweise

Sei Ω ein beschränktes Gebiet und

$$\mathcal{L}u := au_{xx} + 2bu_{xy} + cu_{yy} + du_x + eu_y + qu, \quad \text{für } (x,y) \in \Omega$$

gleichmäßig elliptisch mit $q \leq 0$. Sei $a,b,c,d,e,q \in C(\overline{\Omega})$. Sei $\mathbf{h} = \{h_0, h_1, \dots\}$, wo $h_{k+1} < h_k$ und $h_k \to 0$ für $k \to \infty$. Für $h \in \mathbf{h}$ sei $S_h \subset \overline{\Omega}$ und ℓ_h ein Differenzenoperator auf $R_h \subset S_h$. Sei im weiteren ℓ_h nichtnegativ, R_h ℓ_h-zusammenhängend, $\partial R_h \neq \phi$ und $\partial R_h \subset \partial\Omega$.

$\{\ell_h, R_h, \mathbf{h}\}$ heißt eine gleichmäßig konsistente Approximation zu $\mathcal{L}$ bezüglich $C^{(N)}(\overline{\Omega})$, wenn für alle $u \in C^{(N)}(\overline{\Omega})$

$$\lim_{\substack{h\to 0\\ h\in\mathbf{h}}} \max_{R_h} |\ell_h r_h u - s_h \mathcal{L}u| = 0$$

$\{\ell_h, R_h, \mathbf{h}\}$ heißt eine gleichmäßig konsistente Approximation der Ordnung k zu $\mathcal{L}$ bezüglich $C^{(N)}(\overline{\Omega})$, wenn für alle $u \in C^{(N)}(\overline{\Omega})$

$$\max_{R_h} |\ell_h r_h u - s_h \mathcal{L}u| \leq K h^k \; ; \text{ mit } K \text{ von } h \text{ unabhängig} .$$

Beispiel 3.2 Wenn ℓ_h und $\Omega_h = R_h$ wie in Absatz 4.2, konstruiert werden, dann ist $\{\ell_h, R_h, \mathbf{h}\}$ gleichmäßig konsistent der Ordnung 2 bezüglich $C^{(4)}(\overline{\Omega})$.

Lemma 3.2 *Sei $\{\ell_h, R_h, \mathbf{h}\}$ eine gleichmäßig konsistente Approximation zu $\mathcal{L}$. Dann existiert eine Gitterfunktion ϕ, so daß für $h \in \mathbf{h}$, h genügend klein,*

$$\phi \geq 0 \, , \quad \ell_h \phi \geq 1 \, , \quad \textit{in } R_h \, .$$

Beweis:

Sei $\phi := e^{\alpha x^2}$, wo $\alpha > 0$ eine Konstante ist. Dann ist $\phi \geq 0$. Auch

$$\mathcal{L}\phi = e^{\alpha x^2}[a\alpha^2 + d\alpha + q] \, .$$

Man wähle α, so daß

$$\mathcal{L}\phi \geq 2 \text{ in } \overline{\Omega} \, .$$

Es folgt für h klein

$$|s_h \mathcal{L}\phi - \ell_h r \phi| \leq 1 \text{ in } R_h \, .$$

Satz 3.4 *Sei* $\{\ell_h, R_h, \mathbf{h}\}$ *eine gleichmäßig konsistente Approximation zu* $\mathcal{L}$ *bezüglich* $C^{(N)}(\overline{\Omega})$.

Sei $u \in C^{(N)}(\overline{\Omega})$ *mit*

$$\begin{aligned} \mathcal{L}u &= g \;, \quad \textit{in } (\Omega) \\ u &= \gamma \;, \quad \textit{auf } \partial\Omega \;. \end{aligned}$$

Sei v_h *die Lösung der Differenzengleichungen, deren Existenz und Eindeutigkeit durch Satz 3.2 gewährleistet ist*

$$\begin{aligned} \ell_h v_h &= g \;, \quad \textit{in } R_h \\ v_h &= \gamma \;, \quad \textit{auf } \partial R_h \;. \end{aligned}$$

Dann gilt

$$\lim_{h\to 0} \max_{\overline{R}_h} |v_h(P) - u(P)| = 0 \;.$$

Sei zusätzlich $\{\ell_h, R_h, \mathbf{h}\}$ *eine gleichmäßig konsistente Approximation der Ordnung* k *zu* $\mathcal{L}$ *bezüglich* $C^{(N)}(\overline{R})$, *dann gilt*

$$\max_{\overline{R}_h} |v_h(P) - u(P)| = O(h^k) \;.$$

Beweis:

Sei $w_h := v_h - r_h u$. Aus Lemma 3.2 und Satz 3.3 folgt:

$$\max_{\overline{R}_h} |w_h| \le \max_{\partial R_h} |w_h| + M \max_{R_h} |\ell_h w_h|$$

mit $M = \max_{\partial\Omega} \phi$.
Da $w_h = 0$ auf ∂R_h und $\ell_h v_h = g = \mathcal{L}u$ in R_h, folgt

$$\max_{\overline{R}_h} |w_h| \le M \max_{R_h} |\ell_h r_h u - \mathcal{L}u| \;.$$

Der Satz folgt sofort.

Satz 3.4 stammt von Gerschgorin [1930]. Ein Nachteil dieses Satzes ist, daß vorausgesetzt wird, daß der Differenzenoperator ℓ_h den Differentialoperator $\mathcal{L}$ mit der gleichen Genauigkeit überall in R_h approximiert. Es wurde zuerst von Wasow bemerkt, daß die Ordnung des Fehlers ungeändert ist, wenn ℓ_h $\mathcal{L}$ mit weniger Genauigkeit in der Gegend von $\partial\Omega$ approximiert. Diese Ideen werden jetzt für die Laplace-Gleichung erklärt:

$$\begin{aligned} u_{xx} + u_{yy} &= 0\,, \ \text{in } \Omega \\ u &= \gamma\,, \ \text{auf } \partial\Omega \end{aligned}$$

Sei G_h ein Gitter mit der Maschenweite h und $R_h = G_h \cap \Omega$. Sei

$$w_h := r_h u - v_h$$

der Fehler, wo v_h die Lösung der noch anzugebenden Differenzengleichungen ist.

Sei

$$R_h = R_h^{'} \cup R_h^{''}$$

wo

$$\begin{aligned} R_h^{'} &= \text{Menge der "regulären" inneren Punkte} \\ R_h^{''} &= \text{Menge der "nicht-regulären" inneren Punkte} \end{aligned}$$

d.h. Punkte mit mindestens einem Randpunkt als Nachbar. (Siehe Abbildung 3.12.)

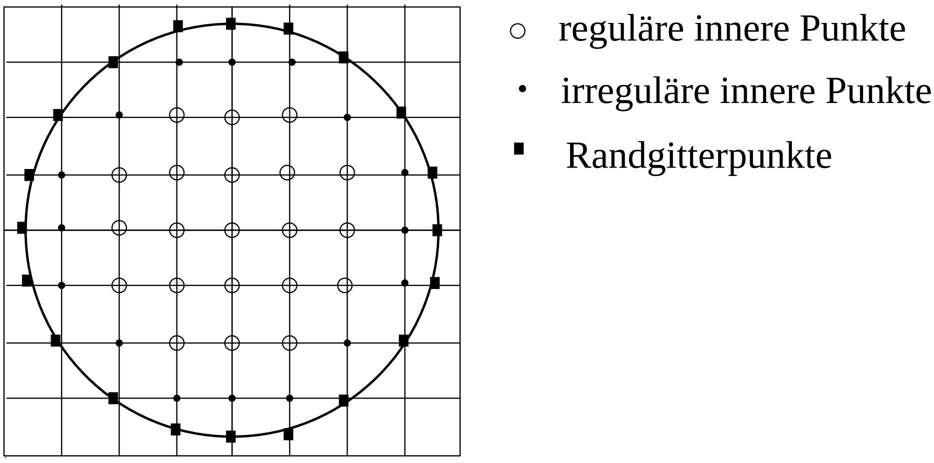

Abbildung 3.12: Reguläre und nicht-reguläre Punkte

Zuerst betrachten wir $R_h^{'}$. Ist $P \in R_h^{'}$, benutzen wir die übliche Fünfpunkt-Formel:

$$\frac{v_h(P_1) + v_h(P_2) + v_h(P_3) + v_h(P_4) - 4v_h(P)}{h^2} = 0$$

mit einem Fehler $0(h^2)$. Folglich - siehe den Beweis von Satz 3.4 - gilt:

$$\|w_h\|_{R_h^{'}} \leq \|w_h\|_{\partial R_h^{'}} + K^{'} h^2 \,.$$

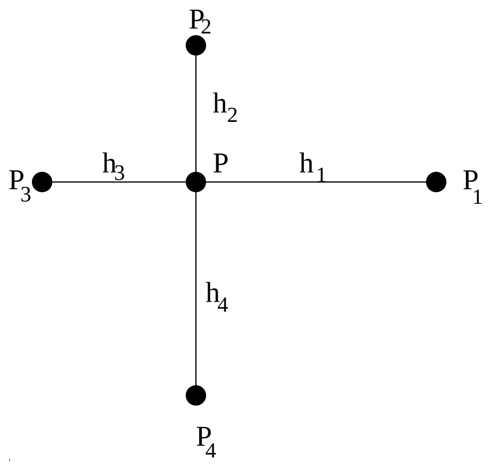

Abbildung 3.13: Die Nachbarn eines nicht-regulären Punktes $P \in R_h^{''}$

Aber $\partial R_h^{'} \subset R_h^{''} \cup \partial\Omega$ und $w_h(P) = 0$ für $p \in \partial R \subset \partial\Omega$, so daß

$$\|w_h\|_{R_h^{'}} \leq \|w_h\|_{R_h^{''}} + K^{'} h^2 \ . \qquad (*)$$

Jetzt betrachten wir $R_h^{''}$. Ist $P \in R_h^{''}$, benutzen wir die Formel (siehe Abbildung 3.13).

$$\begin{aligned}
A_1 v_h(P_1) + A_2 v_h(P_2) + A_3 v_h(P_3) + A_4 v_h(P_4) - A_0 v_h(P_0) &= 0 \\
A_1 &= \frac{2}{h_1(h_1+h_3)} \ , \ A_3 = \frac{2}{h_3(h_1+h_3)} \\
A_2 &= \frac{2}{h_2(h_2+h_4)} \ , \ A_4 = \frac{2}{h_4(h_2+h_4)} \\
A_0 &= \sum_{i=1}^{4} A_i
\end{aligned}$$

mit einem Fehler $0(h)$.

$|w_h|$ erreiche das Maximum in $R_h^{''}$ an der Stelle P_0. Dann gilt

$$h^2 A_0 |w_h(P_0)| \leq \sum_{i=1}^{4} h^2 A_i |w_h(P_i)| + K^{''} h^3 \ ,$$

Es ist leicht einzusehen, daß

1. $h^2 A_i \geq 1$
2. $h^2 A_i \leq 2$, wenn $h_i = h$.
3. Ist $h_i < h$, dann ist $P_i \in \partial R_h$ und $w_h(P_i) = 0$.

Sei

$$\begin{aligned} I^{'} &:= \{i|1 \le i \le 4\,, \quad P_i \in R_h^{'}\} \\ I^{''} &:= \{i|1 \le i \le 4\,, \quad P_i \in R_h^{''}\} \\ I^{'''} &:= \{i|1 \le i \le 4\,, \quad P_i \in \partial R_h\} \end{aligned}$$

$P_0 \in R_h^{''}$, so daß P_0 mindestens einen Randpunkt als Nachbarn hat und $I^{'''}$ nicht leer ist.

Dann ergibt sich

$$h^2 A_0 |w_h(P_0)| \le \|w_h\|_{R_h^{'}} h^2 \sum_{I^{'}} A_i + \|w_h\|_{R_h^{''}} \sum_{I^{''}} A_i h^2 + K^{''} h^3$$

so daß

$$\left(h^2 A_0 - \sum_{I^{''}} h^2 A_i \right) \|w_h\|_{R_h^{''}} \le \sum_{I^{'}} h^2 A_i \|w_h\|_{R_h^{'}} + K^{''} h^3 \,.$$

Da

$$h^2 A_0 - \sum_{I^{''}} h^2 A_i = \sum_{i \in I^{'''}} h^2 A_i + \sum_{i \in I^{'}} h^2 A_i \ge 1 + \sum_{I^{'}} h^2 A_i \,,$$

folgt

$$\|w\|_{R_h^{''}} \le \|w\|_{R_h^{'}} \cdot \frac{a}{a+1} + \frac{K^{''} h^3}{a+1}$$

mit

$$a := \sum_{I^{'}} h^2 A_i \,.$$

Aber $0 \le a \le 6$, so daß

$$\|w\|_{R_h^{''}} \le \frac{6}{7} \|w\|_{R_h^{'}} + K^{''} h^3 \qquad (**)$$

Wir kombinieren die Ungleichungen $(*)$ und $(**)$:

$$\|w\|_{R_h^{'}} \le K^{'} h^2 + \frac{6}{7} \|w\|_{R_h^{'}} + K^{''} h^3 \,,$$

so daß

$$\begin{aligned} \|w\|_{R_h^{'}} &\le 7[K^{'} h^2 + K^{''} h^3] \\ \|w\|_{R_h^{''}} &\le 6[K^{'} h^2 + K^{''} h^3] + K^{''} h^3 \,. \end{aligned}$$

D.h. $\|w_h\| = 0(h^2)$.

Diese Ideen können natürlich für allgemeinere Operatoren ℓ_h angewandt werden. Eine solche Verallgemeinerung befindet sich in Forsythe und Wasow [1960].

Literatur:

Björck, A. und **Dahlquist**, G.: Numerische Methoden. Oldenbourg, München, 1979.

Collatz, L.: Numerische Behandlung von Differentialgleichungen, 2. Auflage. Springer, 1955.

Courant, R., **Friedrichs**, K., **Lewy**, H.: Über die partiellen Differenzengleichungen der mathematischen Physik. Math.Ann. 100, 32-74 (1928).

Forsythe, G.E., **Wasow**, W.R.: Finite Difference Methods for Partial Differential Equations. John Wiley, 1960.

Gerschgorin, S.: Fehlerabschätzung für das Differenzenverfahren zur Lösung partieller Differentialgleichungen. Z. Angew. Math. Mech. 10, 373-382 (1930).

Großmann, Ch., **Roos**, H.-G.: Numerik partieller Differentialgleichungen. Teubner, 1992.

Hackbusch, W.: Theorie und Numerik elliptischer Differentialgleichungen. Teubner, 1986.

Nörlund, N.E.: Vorlesungen über Differenzenrechnungen. Chelsea, New York, 1954. [Neudruck eines Buches aus den zwanziger Jahren.]

Richardson, L.F.: The approximate arithmetical solution by finite differences of physical problems involving differential equations with an application to the stresses in a masonry dam. Phil. Trans. Roy. Soc. London A 210 (308-357), 1911.

Runge, C.: Über eine Methode, die partielle Differentialgleichung Δu = Constans numerisch zu integrieren. Z. Math. u. Phys. 56 (225-232), 1908.

Strikwerda, J.C.: Finite Difference Schemes and Partial Differential Equations. Wadsworth and Brooks/Cole, 1989.

Kapitel 4

Iterationsverfahren zur Lösung großer linearer Gleichungssysteme: klassische Verfahren

4.1 Einleitung

Im vorherigen Kapitel sind Randwertaufgaben für elliptische partielle Differentialgleichungen durch finite Differenzen approximiert worden. Das Problem

$$\begin{aligned} Au &= f \quad \text{auf } \Omega \\ u &= g \quad \text{auf } \partial\Omega \end{aligned} \tag{4.1}$$

wurde dadurch durch das Problem

$$A_h u^h = f_h \tag{4.2}$$

approximiert, wobei

$$A_h : V_h \to F_h \tag{4.3}$$

und V_h und F_h endlichdimensionale Banachräume sind.

Alle anderen numerischen Methoden zur Approximation der Gleichung (4.1) führen ebenfalls zu einer Gleichung der Gestalt (4.2).

Die Gleichung (4.2) hat mehrere Eigenschaften:

1. Die Dimension N des Raumes V_h kann sehr großsein, z.B. $N = 10^6$.

2. Wird die Gleichung (4.2) als eine Matrixgleichung betrachtet, dann ist die Matrix A_h oft dünnbesetzt, symmetrisch, positiv- definit und diagonal dominant. (Dies gilt insbesondere für das Modellproblem $\Delta u = f$.)

Es gilt, diese Eigenschaften von A_h auszunutzen. Zur Lösung der Gleichung (4.2) können entweder iterative oder direkte Methoden angewandt werden. In diesem Abschnitt werden iterative Methoden diskutiert.

Die folgenden Themen werden behandelt:

Das Jacobi Verfahren
Das Gauß-Seidel Verfahren
Mehrgittermethoden
Das konjugierte Gradientenverfahren

4.2 Hilfsmittel

Die folgenden Begriffe und Sätze werden später benötigt (siehe Numerik I):

Definition 4.1 *Sind λ_i , $1 \leq i \leq \ell$, die Eigenwerte von $A \in M(n,n)$, so nennt man*

$$\rho(A) := \max_{1 \leq i \leq \ell} |\lambda_i|$$

den Spektralradius von A .

Definition 4.2 *Eine Matrix $J \in M(r,r)$ heißt Jordanmatrix zum Eigenwert λ , wenn*

$$J = \begin{pmatrix} \lambda & 1 & & O \\ & \lambda & 1 & \\ & & \ddots & 1 \\ O & & & \lambda \end{pmatrix}$$

Satz 4.1 (Jordansche Normalform) *Sei $A \in M(n,n)$. Es gibt eine reguläre Matrix $S \in M(n,n)$, so daß gilt*

$$SAS^{-1} = J = \begin{pmatrix} J_1 & & & \\ & J_2 & & \\ & & \ddots & \\ & & & J_\ell \end{pmatrix}$$

wobei für $1 \leq r \leq \ell$ $J_r \in Mat(n_r, n_r)$ eine Jordanmatrix ist,

$$J_r = \begin{pmatrix} \lambda_r & 1 & & O \\ & \lambda_r & 1 & \\ & & \ddots & 1 \\ O & & & \lambda_r \end{pmatrix}$$

<u>Beweis:</u> Siehe Fischer, Lineare Algebra, Anhang B.

Die charakteristischen Polynome $p_r(\lambda) = \det(J_r - \lambda I) = (\lambda_r - \lambda)^{n_r}$ heißen die <u>Elementarteiler</u> von A .

Satz 4.2 *Für alle Eigenwerte λ von A gilt*

$$|\lambda| \leq lub(A) := \sup_{x \neq 0} \frac{||Ax||}{||x||}$$

für jede Vektornorm $||\cdot||$.

<u>Beweis:</u>
Sei (λ, v) ein Eigenpaar von A , d.h. $Av = \lambda v$ mit $v \neq 0$. Dann folgt

$$|\lambda| \, ||v|| = ||Av|| \leq ||A|| \cdot ||v|| \, .$$

Satz 4.3 *a) Zu jeder Matrix A und jedem $\epsilon > 0$ existiert eine Vektornorm $||\cdot||_\wedge$ mit*

$$lub(A)_\wedge = ||A||_\wedge \leq \rho(A) + \epsilon \, .$$

b) Hat jeder Eigenwert λ von A mit der Eigenschaft $|\lambda| = \rho(A)$ nur lineare Elementarteiler, so existiert sogar eine Vektornorm $||\cdot||_\wedge$ mit

$$||A||_\wedge := lub(A)_\wedge = \rho(A) \, .$$

<u>Beweis von a):</u>
Sei $SAS^{-1} = J$, wo J in Jordannormalform ist, $J = \mathrm{diag}(J_1, \dots, J_\ell)$,

$$J = \begin{pmatrix} J_1 & & \\ & \ddots & \\ & & J_\ell \end{pmatrix}$$

$$J_r = \begin{pmatrix} \lambda_r & 1 & & O \\ & \lambda_r & 1 & \\ & & \ddots & 1 \\ O & & & \lambda_r \end{pmatrix}$$

Sei $C(\epsilon) := C := \mathrm{diag}(C_1, \dots, C_\ell)$ mit

$$C_r = C_r(\epsilon) = \mathrm{diag}(1, \epsilon, \dots, \epsilon^{n_r - 1}) \in \mathrm{Mat}(n_r, n_r) \, .$$

Es folgt

$$C^{-1}JC = \mathrm{diag}(C_1^{-1}J_1C_1, \ldots, C_\ell^{-1}J_\ell C_\ell) \; ,$$

mit

$$C_r^{-1}J_rC_r = \begin{pmatrix} \lambda_r & \epsilon & & O \\ & \lambda_r & \epsilon & \\ & & \ddots & \epsilon \\ O & & & \lambda_r \end{pmatrix}$$

Es folgt sofort, daß

$$\|C^{-1}JC\|_\infty \leq \rho(A) + \epsilon \; ,$$

so daß gilt

$$\|TAT^{-1}\|_\infty \leq \rho(A) + \epsilon$$

mit

$$T := C^{-1}S \; .$$

D.h. (siehe Hilfssatz)

$$\|A\|_\wedge := \|TAT^{-1}\|_\infty \leq \rho(A) + \epsilon \; .$$

Beweis von b):
Für

$$\epsilon < \rho(A) - \max_{|\lambda_r| < \rho(A)} |\lambda_r|$$

setze

$$D_r = \begin{cases} C_r(\epsilon) & , \quad |\lambda_r| < \rho(A) \\ I & , \quad |\lambda_r| = \rho(A) \end{cases}$$

$$D = \mathrm{diag}(D_1, \ldots, D_\ell) \; .$$

Dann gilt:

$$\|D^{-1}JD\|_\infty = \rho(A) \; .$$

Hilfssatz:
Sei $\|\cdot\|$ eine Vektornorm und $T \in \mathrm{Mat}(n,n)$ eine reguläre Matrix. Dann ist $\|\cdot\|_\wedge$,

$$\|x\|_\wedge := \|Tx\|$$

eine Vektornorm und

$$\mathrm{lub}(A)_\wedge = \|TAT^{-1}\| \; , \quad \text{für alle} \quad A \in \mathrm{Mat}(n,n) \; .$$

Beweis:
$\|\cdot\|_\wedge$ ist eine Vektornorm, da die drei notwendigen Voraussetzungen offensichtlich erfüllt sind.

Es gilt für $A \in \mathrm{Mat}(n,n)$,

$$\begin{aligned}
\mathrm{lub}(A)_\wedge &= \sup_{x \neq 0} \frac{\|Ax\|_\wedge}{\|x\|_\wedge} \\
&= \sup_{y \neq 0} \frac{\|AT^{-1}y\|_\wedge}{\|T^{-1}y\|_\wedge} \\
&= \sup_{y \neq 0} \frac{\|TAT^{-1}y\|}{\|y\|} \\
&= \|TAT^{-1}\|
\end{aligned}$$

Satz 4.4 *Sei $A \in Mat(n,n)$. Der Limes*

$$R(A) := \lim_{k \to \infty} \|A^k\|^{1/k}$$

existiert und ist gleich $\rho(A)$.

<u>Beweis:</u>
Sei

$$R_k := \|A^k\|^{1/k}$$

und (λ, x) ein Eigenpaar von A . Dann gilt:

$$R_k \geq \left[\frac{\|A^k x\|}{\|x\|}\right]^{1/k} = |\lambda| \ ,$$

so daß

$$R_k \geq \rho(A) \ . \qquad (*)$$

Sei nun $\epsilon > 0$ beliebig. Es folgt aus Satz 4.3, daß es eine Norm $\|\ \|_\epsilon$ gibt mit

$$\|A\|_\epsilon \leq \rho(A) + \epsilon \ .$$

Alle Normen auf dem $\mathbb{R}^n$ sind äquivalent, es gibt deshalb eine Konstante $m = m(\epsilon)$ mit

$$\frac{1}{m}\|x\| \leq \|x\|_\epsilon \leq m\|x\| \ .$$

Es folgt

$$\begin{aligned}
R_k &= \left[\sup_{x \neq 0} \frac{\|A^k x\|}{\|x\|}\right]^{1/k} \\
&\leq \left[\sup_{x \neq 0} \frac{m^2\|A^k x\|_\epsilon}{\|x\|_\epsilon}\right]^{1/k} \\
&= m^{2/k} \cdot [\|A^k\|_\epsilon]^{1/k} \\
&\leq m^{2/k} \cdot \|A\|_\epsilon \\
&\leq m^{2/k} \cdot (\rho(A) + \epsilon) \qquad (**)
\end{aligned}$$

Für alle $m > 0$ gilt

$$\lim_{k\to\infty} m^{2/k} = 1 .$$

Der Satz folgt deshalb sofort aus den Ungleichungen (*) und (**).

Bemerkung 4.1 Es folgt aus diesem Satz, daß der Spektralradius $\rho(A)$ nicht nur die Konvergenz des Verfahrens $x^{k+1} = Ax + b$, sondern auch dessen asymptotische Konvergenzgeschwindigkeit bestimmt.

Bemerkung 4.2 a) $\rho(A) \leq \|A\|$

b) $\rho(A) < 1$ und $\rho(B) < 1$ impliziert i.a. nicht $\rho(AB) < 1$.

Beispiel 4.1 a)

$$A = \begin{pmatrix} 1/2 & 100 \\ 0 & 1/2 \end{pmatrix} , \quad A^k = \begin{pmatrix} (\frac{1}{2})^k & 100k(\frac{1}{2})^{k-1} \\ 0 & (\frac{1}{2})^k \end{pmatrix}$$

b)

$$A = \begin{pmatrix} 1/2 & 100 \\ 0 & 1/2 \end{pmatrix} , \quad B = \begin{pmatrix} 1/2 & 0 \\ 100 & 1/2 \end{pmatrix}$$

$$AB = \begin{pmatrix} 10^4 + 1/4 & 50 \\ 50 & 1/4 \end{pmatrix}$$

$$\begin{aligned} \rho(A) &= \rho(B) = 1/2 , \\ \rho(AB) &> 1 . \end{aligned}$$

Genauere Information über die Potenzen A^m einer Matrix A erhält man aus der Jordanschen Normalform.

Satz 4.5 *A sei eine $n \times n$-Matrix mit $\rho(A) > 0$. Dann gilt:*

$$\|A^m\| \sim \nu \binom{m}{p-1} [\rho(A)]^{m-(p-1)} , \; m \longrightarrow \infty ,$$

wo p die größte Ordnung aller Jordanscher Blockmatrizen J von A mit $\rho(J) = \rho(A)$ und ν eine positive Konstante ist.

Beweis: Siehe Varga, Matrix Iterative Analysis, S. 65.

Beispiel 4.2

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

$$A^m = \begin{pmatrix} \lambda^m & m\lambda^{m-1} \\ 0 & \lambda^m \end{pmatrix} .$$

4.3 Das Jacobi-, Gauß-Seidel-, SOR- und SSOR-Verfahren - eine Einleitung

In diesem Abschnitt werden vier wichtige Iterationsverfahren zur Lösung des Gleichungssystems

$$Ax = b \tag{4.4}$$

vorgestellt. Im nächsten Abschnitt wird die Konvergenz dieser Verfahren untersucht.

4.3.1 Anwendung auf ein Modellproblem

Als Beispiel nehmen wir

$$A = \begin{pmatrix} 4 & 0 & -1 & -1 \\ 0 & 4 & -1 & -1 \\ -1 & -1 & 4 & 0 \\ -1 & -1 & 0 & 4 \end{pmatrix} , \quad b = \begin{pmatrix} 2 \\ 2 \\ 2 \\ 2 \end{pmatrix} , \quad x = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} \tag{4.5}$$

Die Randwertaufgabe

$$\begin{aligned} \Delta u &= 0 \text{ in } \Omega \\ u &= 1 \text{ auf } \partial\Omega \\ \Omega &:= (0,1) \times (0,1) \end{aligned}$$

werde unter Anwendung des Fünfpunktsterns mit Schrittweite $h = \frac{1}{3}$ approximiert. Die Numerierung der Knoten sei dabei wie in Abbildung 4.1a) .
Es ergibt sich nach Elimination der vorgeschriebenen Randwerte:

$$A_h u_h = f_h$$

$$A_h := \begin{pmatrix} 4 & -1 & -1 & 0 \\ -1 & 4 & 0 & -1 \\ -1 & 0 & 4 & -1 \\ 0 & -1 & -1 & 4 \end{pmatrix}$$

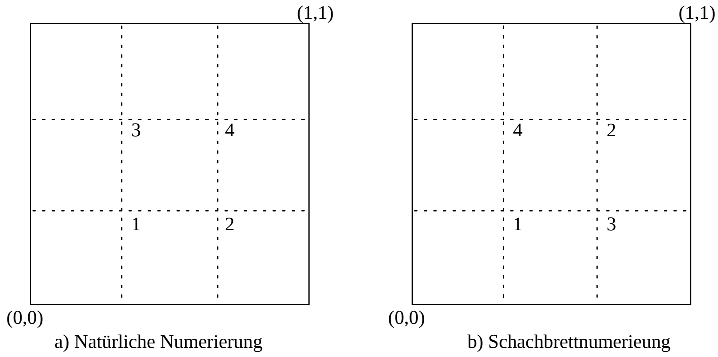

Abbildung 4.1: Mögliche Numerierungen

$$f_h = \begin{pmatrix} g(h,0) & + & g(0,h) \\ g(2h,0) & + & g(1,h) \\ g(h,1) & + & g(0,2h) \\ g(2h,1) & + & g(1,2h) \end{pmatrix} = \begin{pmatrix} 2 \\ 2 \\ 2 \\ 2 \end{pmatrix}$$

$$u_h = \begin{pmatrix} u_1 \\ u_2 \\ u_3 \\ u_4 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}$$

A_h, u_h und f_h werden natürlich durch die Numerierung der Gitterpunkte beeinflußt. Eine Änderung der Numerierung führt dazu, daß A_h, u_h und f_h durch $\tilde{A}_h, \tilde{u}_h$ und $\tilde{f}_h$ ersetzt werden, wobei

$$\tilde{A}_h = P^T A_h P \ , \quad \tilde{u}_h = P^T u_h \ , \quad \tilde{f}_h = P^T f_h$$

und P eine Permutationsmatrix ist.

Für die Numerierung in Abbildung 4.1b) gilt:

$$\tilde{A}_h = \begin{pmatrix} 4 & 0 & -1 & -1 \\ 0 & 4 & -1 & -1 \\ -1 & -1 & 4 & 0 \\ -1 & -1 & 0 & 4 \end{pmatrix} \ , \quad \tilde{f}_h = f_h \ , \quad \tilde{u}_h = u \tag{4.6}$$

mit

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

Bei allen drei Verfahren wird A als Differenz von zwei Matrizen M und N dargestellt,

$$A = M - N \ , \tag{4.7}$$

wobei gilt:

1. M sei regulär
2. M sei "leicht invertierbar", d.h. die Gleichung $My = c$ sei für beliebiges c leicht zu lösen.

Das Gleichungssystem (4.4) kann dann in der Gestalt

$$Mx = b + Nx$$

oder

$$x = M^{-1}b + M^{-1}Nx$$

geschrieben werden. Es liegt deshalb nahe, das Iterationsverfahren

$$x^{(k+1)} = Bx^{(k)} + d \ , \ k \geq 0 \tag{4.8}$$

zu benutzen mit

$$\begin{aligned} B &:= M^{-1}N \ , \\ d &:= M^{-1}b \ . \end{aligned}$$

Es ist manchmal sinnvoll, die folgende Zerlegung von A zu benutzen:

$$\begin{aligned} A &= D - L - R \quad \text{mit} \\ D &= \text{diag}(A) \\ L &= \text{linke Dreiecksmatrix} \\ R &= \text{rechte Dreiecksmatrix} \end{aligned}$$

4.3.2 Das Jacobi-Verfahren oder Gesamtschrittverfahren

Zur Lösung von

$$Ax = b$$

wird die Zerlegung

$$\begin{aligned} A &= M - N \ , \\ M &:= D = \text{diag}(A) \ , \\ N &:= D - A = L + R \end{aligned}$$

benutzt. Es folgt

$$x^{(k+1)} = Bx^{(k)} + d$$

mit

$$\begin{aligned} d &= D^{-1}b \\ B &= D^{-1}(D-A) \\ &= D^{-1}(L+U) \\ &= I - D^{-1}A\,. \end{aligned}$$

Es folgt weiter:

$$Dx^{(k+1)} + (A-D)x^{(k)} = b\,.$$

Dies ist die Gleichung, der die Bezeichnung "Gesamtschrittverfahren" zugrunde liegt. Die Matrix $B = I - D^{-1}A$ wird oft als Jacobi-Matrix bezeichnet.

Beispiel 4.3 Wir betrachten das Modellproblem $Ax = b$ mit $A = \tilde{A}_h,\ b = \tilde{f}_h,\ x = \tilde{u}_h$ aus (4.6). Dann gilt für das Jacobi-Verfahren

$$B = \begin{pmatrix} 0 & 0 & 1/4 & 1/4 \\ 0 & 0 & 1/4 & 1/4 \\ 1/4 & 1/4 & 0 & 0 \\ 1/4 & 1/4 & 0 & 0 \end{pmatrix}.$$

Mit

$$\begin{aligned} x^{(0)} &= (0,0,0,0)^T \quad \text{gilt:} \\ x^{(1)} &= (0,5,0,5,0,5,0,5)^T\,, \\ x^{(2)} &= (0,75,0,75,0,75,0,75)^T \\ x^{(10)} &= (0,999023,0,999023,0,999023,0,999023)^T \\ &\vdots \\ x^{(20)} &= (0,9999,0,9999,0,9999,0,9999)^T \end{aligned}$$

Bemerkung 4.3 Bei Verwendung der Numerierung wie in Abbildung 4.1a) erhält man folgende Jacobimatrix:

$$B = \begin{pmatrix} 0 & 1/4 & 1/4 & 0 \\ 1/4 & 0 & 0 & 1/4 \\ 1/4 & 0 & 0 & 1/4 \\ 0 & 1/4 & 1/4 & 0 \end{pmatrix}$$

4.3.3 Das Gauß-Seidel-(Einzelschritt-) Verfahren

$$\begin{aligned} M &= D - L\,, \\ N &= R\,, \\ B &= G := (D-L)^{-1}R\,, \\ (D-L)x^{k+1} &- Rx^{(k)} = b \end{aligned}$$

Beispiel 4.4 Mit

$$
\begin{array}{rcl}
x^{(0)} & = & (0,0,0,0)^T \quad \text{gilt:} \\
x^{(1)} & = & (0,5, 0,5, 0,75, 0,75)^T \\
x^{(2)} & = & (0,875, 0,875, 0,9375, 0,9375)^T \\
x^{(10)} & = & (0,99998, 0,99998, 0,99999, 0,99999)^T
\end{array}
$$

4.3.4 Das S.O.R.-Verfahren

$$
\begin{array}{rcl}
\omega & \in & (0,2)\ , \\
M & = & \frac{1}{\omega}\, D - L\ , \\
N & = & \frac{1-\omega}{\omega}\, D + R\ , \\
B(\omega) & = & L_\omega := (D-\omega L)^{-1}((1-\omega)D + \omega R)\ , \\
(D-\omega L)x^{(k+1)} & - & ((1-\omega)D + \omega R)x^{(k)} = c
\end{array}
$$

Beispiel 4.5 Mit

$$
\begin{array}{rcl}
x^{(0)} & = & (0,0,0,0)^T \\
\omega & = & 1,5 \quad \text{gilt:} \\
x^{(1)} & = & (0,75, 0,75, 1,3125, 1,3125)^T\ , \\
x^{(10)} & = & (0,99882, 0,99882, 1,00002, 1,00002)^T\ .
\end{array}
$$

Mit

$$
\begin{array}{rcl}
x^{(0)} & = & (0,0,0,0)^T \\
\omega & = & 1,0718 \quad \text{gilt:} \\
x^{(1)} & = & (0,535898, 0,535898, 0,823085, 0,823085)^T \\
x^{(7)} & = & (1,0000, 1,0000, 1,0000, 1,0000)^T
\end{array}
$$

4.3.5 Das SSOR-Verfahren

Dieses Verfahren besteht aus der Zusammenfassung eines Vorwärts- und eines Rückwärts-SOR-Verfahrens.

$$\begin{aligned}
\omega &\in (0,2) \\
M_v &= \frac{1}{\omega}D - L \\
N_v &= \frac{1-\omega}{\omega}D + R \\
M_r &= \frac{1}{\omega}D - R \\
N_r &= \frac{1-\omega}{\omega}D + L \\
S(\omega) &= M_r^{-1}N_r M_v^{-1}N_v \\
d &= M_r^{-1}b + B_r M_v^{-1}b
\end{aligned}$$

Der große Vorteile des SSOR-Verfahrens ist, daß $S(\omega)$ ähnlich zu einer symmetrischen Matrix ist, wenn A symmetrisch ist, so daß dann die Eigenwerte von $S(\omega)$ reell sind.

Mit

$$\begin{aligned}
x^{(0)} &= (0,0,0,0)^T \\
\omega &= 1,5 \quad \text{gilt:} \\
x^{(1)} &= (0,86719, 0,86719, 0,65625, 0,65625)
\end{aligned}$$

$$x^{(10)} = (0,99995, 0,99995, 0,99918, 0,99918)$$

4.4 Allgemeine Konvergenzbetrachtungen

Definition 4.3 *Das Iterationsverfahren*

$$\begin{aligned}
x^{k+1} &= Bx^k + d\ , \\
B &= M^{-1}N\ ,\ d = M^{-1}b
\end{aligned}$$

zur Lösung des Gleichungssystems

$$Ax \equiv (M-N)x = b$$

heißt <u>*konvergent*</u>*, falls für alle Startvektoren* $x^{(0)}$ *und Vektoren* b *die Folge* $\{x^{(k)}\}$ *gegen die exakte Lösung* $x = A^{-1}b$ *konvergiert.*

Wir können jetzt notwendige und hinreichende Bedingungen für Konvergenz herleiten. Der folgende Satz ist sowohl eine Erweiterung (da eine hinreichende Bedingung *gegeben wird) als auch eine Spezialisierung (da nur lineare Abbildungen betrachtet werden) des Banachschen Fixpunktsatzes.*

Satz 4.6 *Das Iterationsverfahren*

$$x^{(k+1)} = Bx^{(k)} + d \qquad (*)$$

ist genau dann konvergent, wenn

$$\rho(B) < 1 \ .$$

Beweis:

a) Sei (*) konvergent. Für jedes d konvergiert die Folge $x^{(0)} := 0$, $x^{(k+1)} = Bx^{(k)} + d$. Sei

$$x := \lim_{k\to\infty} x^{(k)} \ .$$

Es gilt $x = Bx + d$, so daß die Gleichung

$$y = By + d$$

für jedes d eine Lösung hat. D.h. die Abbildung $F := I - B$,

$$F : \mathbb{R}^n \longrightarrow \mathbb{R}^n$$

ist surjektiv und deshalb bijektiv (Fischer, S. 86).

Für den Fehler $f^{(k)} := x - x^{(k)}$ gilt:

$$f^{(k+1)} = Bf^{(k)} \ .$$

Wähle $x^{(0)}$ so, daß $f^{(0)} = v$, wobei $\{\lambda, v\}$ ein Eigenpaar für B ist:

$$\begin{aligned} x^{(0)} &:= 0 \ , \\ d &:= (I - B)v \end{aligned}$$

Es folgt:

$$f^{(k)} = \lambda^k v \ , \ k \in \mathbb{N} \ ,$$

und daher $|\lambda| < 1$.

b) Sei umgekehrt $\rho(B) < 1$.

Es folgt aus Satz 4.3 (§2.2), daß eine Norm existiert mit $\|B\| < 1$. Es folgt, daß die Abbildung $g(u) := Bu + d$ kontrahierend ist. Man benutzt jetzt den Banachschen Fixpunktsatz.

Eine einfache aber oft nützliche notwendige Bedingung für die Konvergenz des Gesamtschrittverfahrens wird jetzt hergeleitet.

Definition 4.4 *Die Matrix $A \in Mat(n,n)$ erfüllt das* <u>*starke Zeilensummenkriterium*</u> *falls*

$$|a_{ii}| > \sum_{\substack{j=1\\ j\neq i}}^{n} |a_{ij}| \ , \ 1 \leq i \leq n \ .$$

Satz 4.7 *Die Matrix A erfülle das starke Zeilensummenkriterium. Dann konvergiert das Gesamtschrittverfahren.*

<u>Beweis:</u>
Es gilt für die Jacobi-Matrix $B = (b_{ij})$,

$$\left.\begin{aligned} \sum_{j=1}^{n} |b_{ij}| &= \sum_{\substack{j=1\\ j\neq i}}^{n} |a_{ij}|/|a_{ii}| < 1 \ , \ 1 \leq i \leq n \ , \\ b_{ii} &= 0 \ , \ 1 \leq i \leq n \end{aligned}\right\} \quad (*)$$

Es gibt jetzt zwei einfache Methoden, um zu zeigen, daß $\rho(B) < 1$.

<u>Methode 1:</u>
Wegen (*) gilt:

$$\rho(B) < \|B\|_\infty < 1 \ .$$

<u>Methode 2:</u>
Sei $\{\lambda, v\}$ ein Eigenpaar von B . Wähle k , so daß

$$|v_k| = \|v\|_\infty = \max_j \ |v_j|$$

und betrachte die k-te Gleichung des Gleichungssystems $\lambda v = Bv$:

$$\begin{aligned} |\lambda v_k| &= \sum_{j=1}^{n} |b_{ij}| \ |v_j| \\ &\leq \|v\|_\infty \sum_{j=1}^{n} |b_{ij}| \\ &< \|v\|_\infty \end{aligned}$$

Es folgt, daß $|\lambda| < 1$.

Satz 4.8 *Sei $A \in Mat(n,n)$, A regulär und L_ω die zugehörige SOR-Iterationsmatrix. Dann gilt:*

$$\rho(L_\omega) \geq |\omega - 1| \ .$$

Beweis:

$$L_\omega = (D - \omega L)^{-1}((1-\omega)D + \omega R) .$$

Es folgt:

$$(D - \omega L)L_\omega = (1-\omega)D + \omega R .$$

Weiter gilt:

$$\begin{aligned}
\det(D) \cdot \det(L_\omega) &= \det(D - \omega L) \cdot \det(L_\omega) \\
&= \det((D - \omega L)L_\omega) , \\
&= \det((1-\omega)D + \omega R) , \\
&= (1-\omega)^n \cdot \det(D) , \quad \text{so daß} \\
\det(L_\omega) &= (1-\omega)^n .
\end{aligned}$$

Das Produkt der evtl. vielfältigen Eigenwerte von L_ω ist gleich dem konstanten Term von $\det(\lambda I - L_\omega)$:

$$\prod_{i=1}^{n} \lambda_i = \det(-L_\omega) = (\omega - 1)^n .$$

Bemerkung 4.4 Es folgt aus Satz 4.8, daß das SOR-Verfahren nur für $\omega \in (0,2)$ konvergent ist.

Satz 4.9 *(Ostrowski-Reich)*
Sei $A = D - E - E^*$ eine $n \times n$ hermitesche Matrix, D eine positiv definite hermitesche Matrix, $D - \omega E$ nichtsingulär für $0 \le \omega \le 2$. Sei

$$L_\omega = (D - \omega E)^{-1}((1-\omega)D + \omega E^*) .$$

a) Wenn A positiv definit ist und $\omega \in (0,2)$, dann ist $\rho(L_\omega) < 1$.

b) Falls E eine linke untere Dreiecksmatrix und D eine Diagonalmatrix ist, gilt auch die Umkehrung.

Beweis:
Zuerst wird eine Hilfsgleichung hergeleitet. Die Fehler $e^m := x - x^m$ erfüllen die Gleichungen:

$$\begin{aligned}
e^{m+1} &= L_\omega e^m && (4.9) \\
(D - \omega E)e^{m+1} &= (\omega E^* + (1-\omega)D)e^m && (4.10)
\end{aligned}$$

Sei

$$\delta^m := e^m - e^{m+1} , \; m \ge 0 . \tag{4.11}$$

Aus (4.10) folgt

$$(D-\omega E)\delta^m = [(D-\omega E)-(\omega E^* + (1-\omega)D)]e^m = \omega A e^m \tag{4.12}$$

und

$$\begin{aligned}
\omega A e^{m+1} &= \omega(D-E-E^*)e^{m+1} \\
&= [(\omega-1)D-\omega E^*]e^{m+1} + (D-\omega E)e^{m+1} \\
&= [(1-\omega)D+\omega E^*](e^m - e^{m+1}) \\
&= (1-\omega)D\delta^m + \omega E^*\delta^m \ .
\end{aligned} \tag{4.13}$$

Werden Gleichungen (4.12) und (4.13) mit ϵ^{m*} bzw. ϵ^{m+1*} multipliziert, wobei $*$ jeweils die konjugierte Transposition bedeutet, dann folgt, daß

$$S_1 := e^{m*}[D-\omega E]\delta^m = \omega e^{m*}Ae^m \quad \text{und} \tag{4.14}$$

$$S_2 := e^{m+1*}[(1-\omega)D+\omega E^*]\delta^m = \omega e^{m+1*}Ae^{m+1} \tag{4.15}$$

Sei

$$S := S_1 - S_2 \ . \tag{4.16}$$

Es gilt:

$$\begin{aligned}
S &= e^{m*}[D-\omega E]\delta^m - e^{m+1*}[(1-\omega)D+\omega E^*]\delta^m \\
&= e^{m*}[D-\omega E]\delta^m - (e^m-\delta^m)^*[(1-\omega)D+\omega E^*]\delta^m \\
&= \delta^{m*}[(1-\omega)D+\omega E^*]\delta^m + e^{m*}[\omega D-\omega E-\omega E^*]\delta^m \\
&= \delta^{m*}[(1-\omega)D+\omega E^*]\delta^m + e^{m*}\omega A\delta^m \\
&= \delta^{m*}[(1-\omega)D+\omega E^*]\delta^m + \delta^{m*}\omega A e^m \ .
\end{aligned} \tag{4.17}$$

Mit Hilfe der Gleichung (4.12) folgt:

$$\begin{aligned}
S &= \delta^{m*}\{[(1-\omega)D+\omega E^*]+[D-\omega E]\}\delta^m \\
&= (2-\omega)\delta^{m*}D\delta^m \ .
\end{aligned} \tag{4.18}$$

Aus (4.14), (4.15), (4.16) und (4.18) folgt:

$$(2-\omega)\delta^{m*}D\delta^m = \omega\{e^{m*}Ae^m - e^{m+1*}Ae^{m+1}\} \ . \tag{4.19}$$

Diese Gleichung ist der Schlüssel zum Konvergenzbeweis, da daraus folgt, daß unter bestimmten Voraussetzungen die Zielfunktion $e^{m*}Ae^m$ monoton fallend ist.

Wähle ein Eigenpaar (λ, v) von L_ω und setze $e^0 = v$. Dann folgt aus (4.19):

$$e^1 = L_\omega e^0 = \lambda e^0 \tag{4.20}$$

$$\delta^0 = (1-\lambda)e^0 \tag{4.21}$$

$$(2-\omega)|1-\lambda|^2 e^{0*}De^0 = \omega(1-|\lambda|^2)e^{0*}Ae^0 \ . \tag{4.22}$$

Wir können jetzt den Beweis durchführen.

a) Sei A positiv definit und $\omega \in (0,2)$. Die Matrix D ist positiv definit. Wegen $e^0 = v \neq 0$ gilt $e^{0*}De^0 > 0$. Es folgt aus (4.22), daß entweder $\lambda = 1$ oder $1-|\lambda|^2 > 0$.

   Wäre $\lambda = 1$, dann folgte aus (4.21), daß $\delta^0 = 0$ und folglich aus (4.12), daß $e^{0*}Ae^0 = 0$, was der Annahme $e^0 = v \neq 0$ widerspricht.

   Es gilt deshalb
$$|\lambda| < 1 .$$
   Da λ einen beliebigen Eigenwert darstellt, ist $\rho(B) < 1$.

b) Sei nun $\rho(L_\omega) < 1$, E eine untere Dreiecksmatrix und D eine Diagonalmatrix. Dann konvergiert die Folge $\{x^k\}$ für jedes x^0, so daß $e^m \to 0$ für $m \to \infty$.

   Es ist bekannt aus Satz 4.8, daß
$$|\omega - 1| \leq \rho(L_\omega) < 1 ,$$
   d.h. $\omega \in (0,2)$.

   Es gilt, da $\omega \neq 0$,
$$A = \frac{1}{\omega}\,(D - \omega E)(I - L_\omega) ,$$
   so daß A regulär ist.

   Wäre A nicht positiv definit, hätte A mindestens einen negativen Eigenwert λ mit Eigenvektor v. Setze $e^0 := v$. Es gelte:
$$e^{0*}Ae^0 < 0 .$$
   Es würde dann aus (4.19) folgen, daß
$$e^{m+1*}Ae^{m+1} \leq e^{m*}Ae^m \leq \cdots \leq e^{0*}Ae^0 < 0 ,$$
   was der Konvergenz der Folge $\{e^m\}$ gegen Null widerspricht.

Zusammenfassend mußgelten:
$$\begin{array}{lcl} \omega & \in & (0,2) , \\ A & & \text{positiv definit} . \end{array}$$

4.5 Die Gerschgorin Sätze

Gerschgorin[1] hat einige Sätze bewiesen, die manchmal bei der Abschätzung von Eigenwerten sehr nützlich sind.

Satz 4.10 (Gerschgorin) *Die Vereinigung aller Kreisscheiben*

$$K_i := \{\mu \in \mathbb{C}| \quad |\mu - a_{ii}| \leq \sum_{\substack{k=1 \\ k \neq i}}^{n} |a_{ij}|\}$$

enthält alle Eigenwerte der $n \times n$-Matrix $A = (a_{ik})$.

Ist die Vereinigung $M_1 = \bigcup_{i=1}^{k} K_i$ disjunkt von der Vereinigung M_2 der übrigen Kreise, so enthält M_1 genau k Eigenwerte von A und M_2 genau $n-k$ Eigenwerte.

Beweis: Stoer und Bulirsch II, S. 77 und 78.

Für das Modellproblem 1 haben die vier Gerschgorin Kreise alle den Mittelpunkt $(0,0)$ und den Radius $\frac{1}{2}$, so daß $\rho(J) \leq \frac{1}{2}$.

Dieses Beispiel ist aber eine Ausnahme, weil keine echten inneren Gitterpunkte existieren. Sobald alle fünf Gitterpunkte eines Sterns innere Gitterpunkte sind, liefert Satz 4.10 nur die Abschätzung $\rho(J) \leq 1$, was die Konvergenz des Gesamtschrittverfahrens nicht garantiert. Es ist deshalb notwendig, einige weitere Begriffe einzuführen.

Eine $n \times n$-Matrix A heißt unzerlegbar (irreduzibel), falls es keine Permutationsmatrix P gibt, so daß P^TAP die Gestalt

$$P^TAP = \begin{bmatrix} \tilde{A}_{11} & \tilde{A}_{12} \\ 0 & \tilde{A}_{22} \end{bmatrix}$$

besitzt, wo $\tilde{A}_{11}$ eine $p \times p$-Matrix, $\tilde{A}_{22}$ eine $q \times q$-Matrix mit $p+q=n$, $p>0$, $q>0$ ist.

Die Unzerlegbarkeit einer Matrix A kann man häufig leicht mit Hilfe des der Matrix A zugeordneten (gerichteten) Graphen $G(A)$ prüfen. Wenn A eine $n \times n$-Matrix ist, so besteht $G(A)$ aus n Knoten $P_1, \ldots, P_n$, und es gibt eine gerichtete Kante $P_i \to P_j$ in $G(A)$ genau dann, wenn $a_{ij} \neq 0$.
Z.B. ist der Graph von

$$A = \begin{bmatrix} 1 & 2 & 0 \\ -1 & 1 & 0 \\ 3 & 0 & 1 \end{bmatrix}$$

[1]Gerschgorin, S.: Über die Abgrenzung der Eigenwerte einer Matrix. Izv. Akad. Nauk SSSR Ser.Mat. 7, 749-754 (1931)

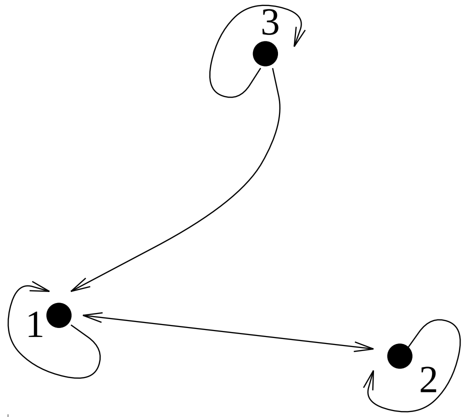

Abbildung 4.2: Der Graph $G(A)$

in Abbildung 4.2 dargestellt.

A ist genau dann unzerlegbar, wenn der Graph $G(A)$ in dem Sinne zusammenhängend ist, daß es für jedes Knotenpaar (P_i, P_j) in $G(A)$ einen gerichteten Weg von P_i nach P_j gibt.

Die Matrix A ist also genau dann unzerlegbar, wenn es für jedes i, j mit $1 \leq i\ ,\ j \leq n$ eine Folge $i_0 i_1 \cdots i_m$ gibt mit

$$\begin{aligned} i_0 &= i \\ i_m &= j \\ a_{i_k i_{k+1}} &\neq 0 \end{aligned}$$

für $0 \leq k < m$.

Behauptung 4.1 Die Matrix A_h entstehe durch die Benutzung des Fünfpunktesystems auf dem Gitter G_h. Dann ist die Matrix A_h genau dann unzerlegbar, wenn G_h zusammenhängend ist.

Satz 4.11 *$A \in Mat(n \times n, \mathbb{C})$ sei unzerlegbar. Sei λ Eigenwert von A, $\lambda \in \partial K$, wobei K die Vereinigung aller Gerschgorin Kreisscheiben ist. Dann gilt:*

$$\lambda \in \partial K_i\ ,\ 1 \leq i \leq n\ .$$

<u>Beweis:</u>

Sei (λ, x) ein Eigenpaar von A mit $\lambda \in \partial K$. O.E. gibt es r mit

$$|x_r| = 1 \geq |x_i|\ ,\ 1 \leq i \leq n\ .$$

Aus

$$Ax = \lambda x$$

folgt

$$\sum_{i=1}^{n} a_{ri} x_i = \lambda x_r$$

und

$$\begin{aligned} |a_{rr} - \lambda| &\leq \sum_{\substack{j=1\\ j\neq r}} |a_{rj}| \cdot |x_j| \\ &\leq \sum_{\substack{j=1\\ j\neq r}} |a_{rj}| \\ &= \text{Radius von} \quad K_r \ , \end{aligned}$$

also $\lambda \in K_r$. Die Voraussetzung, daß $\lambda \in \partial K$, führt dazu, daß $\lambda \in \partial K_r$ und

$$|a_{rj}| \neq 0 \Rightarrow |x_j| = 1 \ , \ 1 \leq j \leq n \ .$$

Da A unzerlegbar ist, gibt es ein $p \neq r$ mit $a_{rp} \neq 0$, so daß $|x_p| = 1$. Durch Wiederholung ergibt sich:

$$|x_i| = 1 \quad \text{für} \quad 1 \leq i \leq n \ ,$$

und $\lambda \in \partial K_i$, für $1 \leq i \leq n$.

Durch Anwendung von Satz 4.11 erhält man:

Satz 4.12 (Schwaches Zeilensummenkriterium) *Falls A unzerlegbar ist und*

$$|a_{ii}| \geq \sum_{k\neq i} |a_{ik}| \quad \textit{für alle} \ \ i = 1, 2, \ldots, n \ ,$$

aber $|a_{i_{0,0}}| > \sum_{k\neq i_0} |a_{i_0 k}|$ für mindestens ein i_0 gilt, dann konvergiert das Gesamtschrittverfahren.

Bemerkung 4.5 Die Jacobi Matrix J ist symmetrisch. Falls J zerlegbar ist, können die einzelnen Untermatrizen $J_1, \ldots, J_k$ getrennt betrachtet werden.

Satz 4.13 *Das Gesamtschrittverfahren konvergiert für die Fünfpunkt- Differenzenapproximation zur Lösung des Dirichlet Problems für die Laplace Gleichung.*

<u>Beweis:</u> Siehe Bemerkung oben.

4.6 Das Gesamtschrittverfahren - das Modellproblem

Es wurde im letzten Abschnitt bewiesen, daß das Gesamtschrittverfahren für das Dirichlet Problem konvergiert. Es stellt sich aber heraus, daß die Konvergenzgeschwindigkeit viel zu wünschen übrig läßt, wie das folgende Modellproblem zeigt.

Beispiel 4.6 Zu betrachten ist das Problem:

$$\begin{aligned} -\Delta u &= f(x,y)\ ,\ \text{in}\ \ \Omega \\ u &= 0\ ,\ \text{auf}\ \ \partial\Omega\ , \end{aligned}$$

wobei

$$\begin{aligned} \Omega &:= (0,1)\times(0,1) \\ f &\in C(\Omega)\ . \end{aligned}$$

Mit der Maschenweite

$$h = \frac{1}{N+1}$$

entsteht das Differenzengleichungssystem

$$A_h v_h = f_h\ ,$$

wobei A_h eine Blocktridiagonalmatrix ist:

$$A_h = \frac{1}{h^2}\begin{pmatrix} H & C & & & \\ C & H & \ddots & O & \\ & \ddots & \ddots & \ddots & \\ & O & \ddots & H & C \\ & & & C & H \end{pmatrix}$$

$H, C \in Mat(N, N, \mathbb{R})$:

$$H = \begin{pmatrix} 4 & -1 & & & \\ -1 & 4 & \ddots & O & \\ & \ddots & \ddots & \ddots & \\ & O & \ddots & 4 & -1 \\ & & & -1 & 4 \end{pmatrix}\ ,$$

$C = -I_N = -I.$

Die entsprechende Jacobi Matrix B ist:

$$B = \frac{1}{4}\begin{pmatrix} F & I & & & \\ I & F & \ddots & O & \\ & \ddots & \ddots & \ddots & \\ & O & \ddots & F & I \\ & & & I & F \end{pmatrix},$$

$$F = \begin{pmatrix} 0 & 1 & & & \\ 1 & 0 & \ddots & O & \\ & \ddots & \ddots & \ddots & \\ & O & \ddots & 0 & 1 \\ & & & 1 & 0 \end{pmatrix};$$

mit $B \in Mat(M, M, \mathbb{R}),\ F \in Mat(N, N, \mathbb{R}),\ M := N^2$.

Das Eigenwertproblem

$$Bv = \lambda v$$

ist äquivalent zum Gleichungssystem:

$$\frac{1}{4}[v_{i,j+1} + v_{i,j-1} + v_{i+1,j} + v_{i-1,j}] = \lambda v_{i,j}\ ,\ \ 1 \leq i, j \leq N\ ,$$

wobei

$$v_{i,j} := 0 \quad \text{für}\ \ i = 0\ ,\ i = N+1\ ,\ j = 0\ ,\ j = N+1\ .$$

Der Ansatz

$$v_{i,j}^{k\ell} = \sin(k\pi hi)\sin(\ell\pi hj)$$

führt zu dem Eigenwert:

$$\begin{aligned} \lambda^{k,\ell} &= \frac{1}{2}[\cos\ k\pi h + \cos\ \ell\pi h] \\ &= 1 - \left(\sin\frac{k\pi h}{2}\right)^2 - \left(\sin\frac{\ell\pi h}{2}\right)^2 . \end{aligned}$$

Die N^2 Vektoren $v^{k,\ell}$ sind linear unabhängig und bilden eine Basis des $\mathbb{R}^{N^2}$. Weiter gilt:

$$\rho(B) = 1 - 2\ \sin^2\ \frac{\pi h}{2} = \cos\ \pi h\ .$$

Wird das Gesamtschrittverfahren zur Lösung des Gleichungssystems

$$A_h v_h = f_h$$

benutzt, dann folgt mit Hilfe von Satz 4.13 aus dem vorigen Abschnitt, daß

$$\sup_{v_h^0 \neq v_h} \limsup_{k\to\infty} \left[\frac{\|v_h^{(k)} - v_h\|}{\|v_h^{(0)} - v_h\|} \right]^{1/k} = \rho(B) \ ,$$

d.h. im Durchschnitt und für allgemeine Anfangsdaten $v_h^{(0)}$ wird der Fehler $\|v_h^{(k)} - v_h^{(0)}\|$ mit dem Faktor $\rho(B)$ pro Iteration multipliziert. Damit sich der Fehler um einen Faktor $1/e$ verringert, werden im Durchschnitt m Iterationen benötigt, wobei

$$\begin{aligned} [\rho(B)]^m &= 1/e \\ m &= -1/\ln(\rho(B)) \\ &\doteq \frac{2}{\pi^2 h^2} \ . \end{aligned}$$

Für großes N (kleines h) sind deshalb sehr viele Iterationen nötig. Der durchschnittliche Rechenaufwand, um den Fehler um einen Faktor $1/e$ zu reduzieren, ist

$$O\left(\frac{2}{\pi^2 h^2}\right) \cdot \text{Multiplikationen} \quad + O\left(\frac{2}{\pi^2 h^2} \cdot 3\right) \ \text{Additionen} \ = O\left(\frac{8}{\pi^2} N^2\right) \ \text{Operationen} \ .$$

Die vorherigen Überlegungen können von einem anderen Gesichtspunkt aus betrachtet werden. Für die Gleichung

$$A_h v_h = f_h$$

ist das Gesamtschrittverfahren

$$D v^{k+1} + (A_h - D) v^k = f_h$$

oder

$$v^{k+1} = B v^k + D^{-1} f_h$$

mit

$$D := \frac{1}{h^2} \mathrm{diag}(-4) \ , \quad B = I - \frac{h^2}{4} A_h \ .$$

Die Fehler

$$e^k := v_h - v^k$$

erfüllen die Gleichung

$$e^{k+1} = Be^k \ .$$

Seien nun (λ_i, w_i) , $1 \le i \le M$, M Eigenpaare von B mit $|\lambda_1| \ge |\lambda_2| \ge \ldots \ge |\lambda_M|$. Der Fehler e^0 läßt sich folgendermaßen darstellen:

$$e^0 = \sum_{i=1}^{M} a_i w_i \ ,$$

so daß

$$e^k = \sum_{i=1}^{M} a_i \lambda_i^k w_i \ .$$

Es folgt - wie bei der Vektoriteration - daß der Fehler für großes k sich etwa wie $a_1 \lambda_1^k w_1$ verhält,

$$e^k \sim a_1 \lambda_1^k w_1 \ .$$

Die asymptotische Konvergenzgeschwindigkeit wird deshalb durch

$$|\lambda_1| = \rho(B)$$

bestimmt. Für kleines k dagegen ist im allgemeinen die Konvergenzgeschwindigkeit etwas besser, da die Fehlerkomponente $a_2 w_2, \ldots, a_M w_M$ sich schneller abbauen läßt. Gerade diese Eigenschaft wird bei Mehrgitterverfahren genutzt.

Weiter gilt:

$$\begin{aligned} \lambda_1 &= \cos \pi h = 1 - 2\sin^2 \frac{\pi h}{2} \ , \\ w_1 &= v_{i,j}^{1,1} = \sin(\pi h i)\sin(\pi h j) \ , \end{aligned}$$

so daß w_1 eine nichtnegative Gitterfunktion ist.

Zusammenfassend gelten die folgenden Aussagen für das Modellproblem, aber auch für allgemeinere nichtnegative Differenzenapproximationen:

1. Das Jacobi-Verfahren konvergiert.

2. $\rho(B) \approx 1 - h^2 \mu_1$, mit $\mu_1 > 0$.

3. Der Fehler nach k Schritten

$$e^k = v_h - v_h^{(k)}$$

läßt sich gut durch

$$e^k \approx \text{constant } \rho(B)^k \cdot w_1$$

approximieren, wobei $w_1 > 0$ eine "glatte" Gitterfunktion ist.

4. Die Anzahl der Iterationen, die benötigt werden, um den Fehler durchschnittlich um einen Faktor $1/e$ zu verringern, ist:

$$m = -1/\ln(\rho(B)) \approx \frac{1}{h^2\mu_1} \, .$$

4.7 Verallgemeinerungen

Die Aussagen im letzten Abschnitt für das Modellproblem sind richtungsweisend für allgemeine Probleme.

Die Tatsache, daß $\rho(B) \approx 1$ für die Jacobi Matrix B gilt, läßt sich folgendermaßen erklären:

Die Eigenwerte μ und Eigenfunktionen u einer schwingenden Membran mit Oberfläche Ω erfüllen die Bedingungen

$$\begin{aligned} \Delta u &= \mu u \, , \text{ in } \Omega \\ u &= 0 \, , \text{ auf } \partial\Omega \, . \end{aligned}$$

Bekanntlich gibt es unendlich viele Eigenpaare $\{\mu_i, u_i\}$ mit $0 < \mu_1 \leq \mu_2 \leq \dots$. Weiter gilt: $\mu_i \to \infty$ für $i \to \infty$; Span$\{u_i\} = L_2(\Omega)$ und $\{u_i\}$ ist eine orthogonale Basis für $L_2(\Omega)$. (Michlin [1978, S. 372]; Garabedian [1964, Chapter 11]; Dunford und Schwartz [1963, S. 1743].)

Das diskrete Eigenwertproblem

$$A_h w_h = \mu_h w_h$$

ist eine Approximation für das Membranproblem. Man erwartet, daß die Eigenwerte $\mu_h^{(i)}$ die Eigenwerte $\mu^{(i)}$ approximieren sollen, d.h.

$$\mu_h^{(i)} \approx \mu^{(i)} .$$

Es folgt, da $B = I - \frac{h^2}{4} A_h$, daß

$$\begin{aligned} \lambda_i &= 1 - \frac{h^2}{4}\, \mu_h^{(i)} \\ &\approx 1 - \frac{h^2}{4}\, \mu^{(i)} . \end{aligned}$$

Für das Rechteck $\Omega = (0,1)^2$ ist

$$\begin{aligned} \lambda &= (k^2 + \ell^2)\pi^2 \quad \text{und} \\ u &= \sin(k\pi x)\sin(\ell\pi y) \end{aligned}$$

für $k, \ell \in \mathbb{N}$ ein Eigenpaar für die schwingende Membran. Es folgt

$$\lambda_1 \approx 1 - \frac{h^2\pi^2}{4}(1^2 + 1^2) = 1 - \frac{h^2\pi^2}{2} ,$$

was das exakte Ergebnis

$$\rho(B) = 1 - 2\ \sin^2\ \frac{\pi h}{2}$$

gut widerspiegelt.

Die Berechnung der Eigenwerte von A_h ist eine interessante Aufgabe. Siehe z.B. Cullum und Willoughby [1985].

4.8 Nichtnegative Matrizen

Für das Modellbeispiel ist $A = D - L - R$ und $B = L + R$, wobei B nichtnegativ und unzerlegbar ist. Es gibt eine schöne Theorie der nichtnegativen Matrizen, die viele Anwendungen hat und leider hier nur kurz aufgegriffen werden kann. (Siehe: Varga [1962], Berman und Plemmons [1979], Seneta [1973], Minc [1988], Windisch [1989].)

Folgender schöner Satz von Perron und Frobenius wurde zwischen 1907 und 1912 bewiesen:

Satz 4.14 (Perron, Frobenius) *Sei A eine $n \times n$ unzerlegbare Matrix, die nichtnegativ ist, d.h. die nichtnegative Komponenten hat. Dann gilt:*

1. *Es gibt einen Eigenwert* λ_+ *von* A *mit* $\lambda_+ = \rho(A) > 0$.

2. *Zu* λ_+ *gehört ein Eigenvektor* $w_+ > 0$.

3. λ_+ *ist ein einfacher Eigenwert von* A.

4. *Sei* $B \geq A$ *und* $B \neq A$, *dann gilt* $\rho(B) > \rho(A)$.

Beweis: Siehe z.B. Varga [1962, S. 30].

Satz 4.15 (Perron, Frobenius - Erweiterung) *Sei* A *eine* $n \times n$ *nichtnegative Matrix. Dann gilt:*

1. *Es gibt einen Eigenwert* λ_+ *von* A *mit* $\lambda_+ = \rho(A) \geq 0$.

2. *Sei* $\lambda_+ = 0$. *Dann ist* A *zerlegbar und die Jordansche Normalform von* A *ist eine echte rechte Dreiecksmatrix.*

3. *Zu* λ_+ *gehört ein Eigenvektor* $w_+ \geq 0$.

4. *Sei* $B \geq A$, *dann gilt* $\rho(B) \geq \rho(A)$.

Beweis: Siehe Varga [1962, S. 46].

Sei

$$\begin{aligned} A &= D - L - R \\ B &:= D^{-1}(L + R) \\ \mathcal{L}_\omega &:= (D - \omega L)^{-1}[(1 - \omega)D + \omega U] \,, \end{aligned}$$

so daß B die Jacobi Matrix ist, $\mathcal{L}_\omega$ die SOR Iterationsmatrix und $\mathcal{L}_1$ die Gauß-Seidel Iterationsmatrix.

Es kann durch Beispiele gezeigt werden, daß sowohl $\rho(B) < 1 < \rho(\mathcal{L}_1)$ als auch $\rho(\mathcal{L}_1) < 1 < \rho(B)$ möglich ist. Unter zusätzlichen Bedingungen erhält man allerdings mit Hilfe des Perron-Frobenius-Satzes den folgenden Satz:

Satz 4.16 (Stein, Rosenberg) *Sei die Matrix* $B = L + R$ *nichtnegativ (*L *und* R *echte linke bzw. rechte Dreiecksmatrizen), dann gilt für* B *und* $\mathcal{L}_1 = (I - L)^{-1}U$ *genau eine der folgenden Beziehungen:*

1. $\rho(\mathcal{L}_1) = \rho(B) = 0$

2. $0 < \rho(\mathcal{L}_1) < \rho(B) < 1$

3. $\rho(\mathcal{L}_1) = \rho(B) = 1$

4. $1 < \rho(B) < \rho(\mathcal{L}_1)$

Beweis: Siehe Varga, S. 70.

4.9 Das SOR Verfahren

Der Satz von Stein und Rosenberg besagt, daß unter den gegebenen Voraussetzungen das Gauß-Seidel-Verfahren besser als das Jacobi- Verfahren ist.

Es ist i.a. nicht möglich, eine Beziehung zwischen $\rho(B)$ und $\mathcal{L}_1$ zu finden, wie das folgende Beispiel von Young zeigt. Für das Modellproblem mit $h = 1/3$ und nur vier inneren Gitterpunkten werden die Gitterpunkte auf drei verschiedene Weisen numeriert:

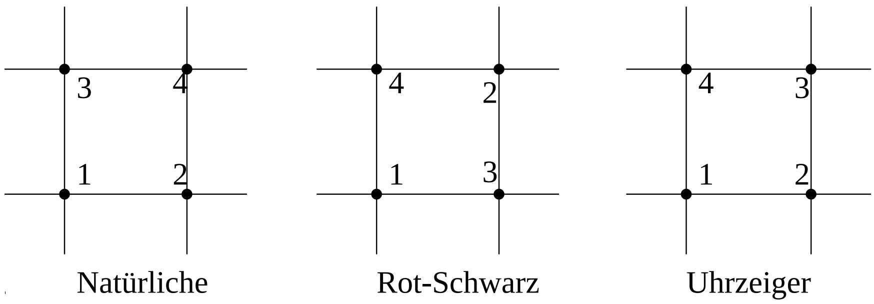

Abbildung 4.3: Verschiedene Numerierungen

Mit der Bezeichnung $\Lambda(A)$ = Spektrum von A = Menge aller Eigenwerte von A folgt:

1. Natürliche Numerierung

$$B = \frac{1}{4} \cdot \begin{bmatrix} 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \end{bmatrix} \qquad \mathcal{L}_1 = \frac{1}{8} \cdot \begin{bmatrix} 0 & 8 & 8 & 0 \\ 0 & 2 & 2 & 8 \\ 0 & 2 & 2 & 8 \\ 0 & 1 & 1 & 1 \end{bmatrix}$$

$\rho(B) = \frac{1}{2}$ $\qquad$ $\rho(\mathcal{L}_1) = \frac{1}{4}$
$\Lambda(B) = \{-1/2, 0, 0, 1/2\}$ $\qquad$ $\Lambda(\mathcal{L}_1) = \{0, 0, 0, 1/4\}$

2. Rot-Schwarz-Numerierung

$$B = \frac{1}{4} \cdot \begin{bmatrix} 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \end{bmatrix} \qquad \mathcal{L}_1 = \frac{1}{8} \cdot \begin{bmatrix} 0 & 0 & 2 & 2 \\ 0 & 0 & 2 & 2 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 \end{bmatrix}$$

$\rho(B) = \frac{1}{2}$ $\qquad\qquad$ $\rho(\mathcal{L}_1) = \frac{1}{4}$
$\Lambda(B) = \{-1/2, 0, 0, 1/2\}$ $\qquad$ $\Lambda(\mathcal{L}_1) = \{0, 0, 0, 1/4\}$

3. Uhrzeiger-Numerierung

$$B = \frac{1}{4} \cdot \begin{bmatrix} 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \end{bmatrix} \qquad \mathcal{L}_1 = \frac{1}{256} \cdot \begin{bmatrix} 0 & 64 & 0 & 64 \\ 0 & 16 & 64 & 16 \\ 0 & 4 & 16 & 68 \\ 0 & 17 & 4 & 33 \end{bmatrix}$$

$\rho(B) = \frac{1}{2}$ $\qquad\qquad$ $\Lambda(\mathcal{L}_1) = \{0, 0, 277, -0,011 \pm 0,118i\}$
$\Lambda(B) = \{-1/2, 0, 0, 1/2\}$

Unter geeigneten Voraussetzungen ist es allerdings möglich, $\rho(\mathcal{L}_\omega)$ genauer zu beschreiben.

Die folgenden Begriffe sind nützlich:

Definition 4.5 *Die Matrix A besitzt die "property A", wenn es eine Permutationsmatrix P gibt, so daß PAP^T die Gestalt*

$$PAP^T = \begin{pmatrix} D_1 & M_1 \\ M_2 & D_2 \end{pmatrix}, \quad D_1, D_2 \text{ Diagonalmatrizen}$$

besitzt.

Definition 4.6 *Die Matrix $A = D(I - L - R)$, wobei L und R echte linke bzw. rechte Dreiecksmatrizen sind, heißt "konsistent geordnet", falls die Eigenwerte von*

$$\alpha L + \alpha^{-1} R\,, \quad \alpha \neq 0\,,$$

von α unabhängig sind.

Bemerkung 4.6 Property $A \Rightarrow$ konsistent geordnet.

Satz 4.17 *Sei*

$$B = \begin{pmatrix} O & B_1 \\ B_2 & O \end{pmatrix}.$$

Sei μ ein Eigenwert von B zum Eigenvektor $(u, v)^T$.

Dann ist $-\mu$ ein Eigenwert von B zum Eigenvektor $(u, -v)^T$.

Beweis: Siehe Birkhoff und Lynch [1984, S. 133].

Satz 4.18 *$(u,v)^T$ sei ein Eigenvektor der Jacobi Matrix B zu A, wobei B die Gestalt*

$$B = \begin{pmatrix} 0 & B_1 \\ B_2 & 0 \end{pmatrix}$$

habe, mit Eigenwert $\mu \neq 0$. Dann gilt:

a) $(u,0)^T$ ist ein Eigenvektor der Gauß-Seidel Matrix $\mathcal{L}_1$ mit Eigenwert 0 und

b) $(u,\mu v)^T$ ist ein Eigenvektor der Gauß-Seidel Matrix mit Eigenwert μ^2.

Weiter gilt:

c) Kern$(\mathcal{L}_1) \supset$ Kern(B).

Beweis: (Siehe Birkhoff, Lynch [1984, S. 134].)

Es folgt, daß für Matrizen, die Property A besitzen, das Gauß-Seidel Verfahren zweimal so schnell wie das Jacobi-Verfahren ist. Es entsteht die Frage, ob es möglich ist, ein ω_b zu finden mit

$$\rho(\mathcal{L}_{\omega_b}) = \inf_{\omega} \rho(\mathcal{L}_\omega) \ .$$

Wir betrachten jetzt symmetrische Matrizen A mit Property A. Für die Jacobi Matrix B gilt dann:

Falls μ ein Eigenwert ist, so ist auch $-\mu$ ein Eigenwert. Die Eigenwerte von $\mathcal{L}_1$ verteilen sich folgendermaßen:

a) Es gibt eine Menge von Eigenwerten 0.

b) Die übrigen Eigenwerte sind positiv und liegen zwischen 0 und $\rho(B)^2$.

Es kann bewiesen werden, daß sich die Eigenwerte von $\mathcal{L}_\omega$ für zunehmendes ω wie in Abbildung 4.4 verhalten: die Nulleigenwerte von $\mathcal{L}_1$ werden größer, die positiven Eigenwerte von $\mathcal{L}_1$ werden kleiner. Wenn zwei Eigenwerte sich treffen, bewegen sie sich weiter in der komplexen Ebene auf dem Kreis $|z| = \omega - 1$.
Das Verhältnis von $\rho(\mathcal{L}_\omega)$ zu ω wird in Abbildung 4.5 dargestellt.
Das SOR Verfahren erreicht den minimalen Wert von $\rho(\mathcal{L}_\omega)$ bei

$$\omega = \omega_b := 1 + \frac{1 - \sqrt{1-\rho(B)^2}}{1 + \sqrt{1-\rho(B)^2}} \ .$$

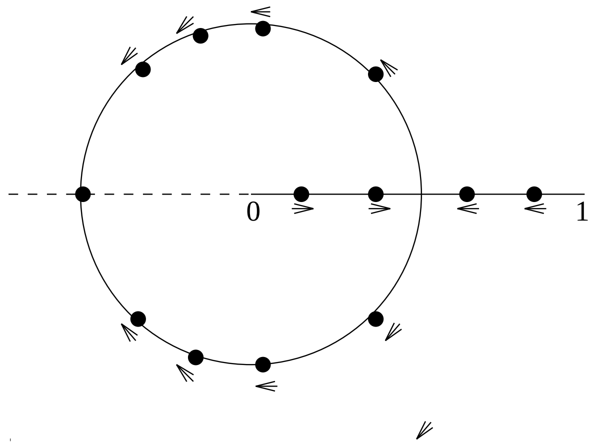

Abbildung 4.4: Die Bewegung der Eigenwerte von $\mathcal{L}_\omega$ mit zunehmendem ω

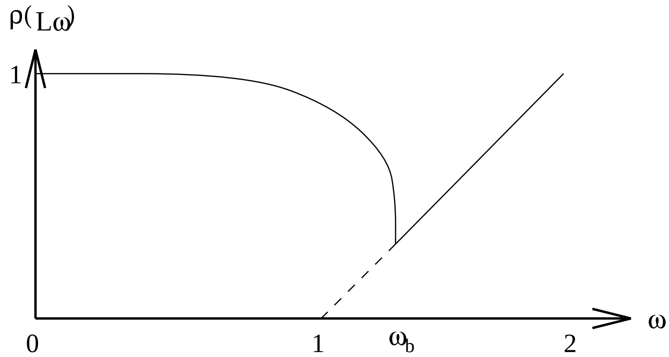

Abbildung 4.5: Die Funktion $\rho(\mathcal{L}_\omega)$

Satz 4.19 (Young, Varga) *A sei eine konsistent geordnete Matrix. Ferner seien die Eigenwerte von B reell und es gelte $\mu = \rho(B) < 1$. Dann ist*

$$\omega_b = \frac{2}{1+\sqrt{1-\rho(B)^2}}\ , \quad \rho(\mathcal{L}_{\omega_b}) = \omega_b - 1 = \left(\frac{\rho(B)}{1+\sqrt{1-\rho(B)^2}}\right)^2 .$$

Allgemein gilt

$$\rho(\mathcal{L}_\omega) = \begin{cases} \omega - 1 & \textit{für } \omega_b \le \omega \le 2 \\ 1-\omega+\frac{1}{2}\omega^2\mu^2+\omega\mu\sqrt{1-\omega+\frac{1}{4}\omega^2\mu^2} & \textit{für } 0 \le \omega \le \omega_b \end{cases}$$

Beweis: Siehe z.B. Stoer und Bulirsch, S. 256.

Für das Modellproblem gilt:

$$\begin{aligned} \rho(B) &\approx 1-\pi^2h^2/2 \\ \rho(\mathcal{L}_1) &\approx 1-\pi^2h^2 \\ \rho(\mathcal{L}_\omega) &\approx 1-2\pi h \end{aligned}$$

Die Geschwindigkeitsverbesserung, die mit SOR erreichbar ist, ist aus Tabelle 4.1 ersichtlich, wobei $N(\mathcal{L}_1)$ die Anzahl der Iterationen ist, die nötig sind, um den Fehler um einen Faktor 1000 zu vermindern.

$1/h$	10	100	200
$\rho(\mathcal{L}_1)$	0,90451	0,99901	0,99975
ρ_{opt} (SOR)	0,52786	0,93909	0,96907
$N(\mathcal{L}_1)$	69	6998	27995
N_{opt} (SOR)	17	195	413

Tabelle 4.1: Vergleich der Gauß-Seidel und SOR Verfahren für das Modellproblem

Literatur:

Berman, A., **Plemmons**, R.J.: Nonnegative Matrices in the Mathematical Sciences. Academic Press, 1979.

Birkhoff, G., **Lynch**, R.E.: Numerical Solution of Elliptic Problems. SIAM 1984.

Cullum, J., **Willoughby**, R.A.: Lanczos Algorithms for Large Symmetric Eigenvalue Problems. Birkhäuser, 1985. (Programme sind mit Hilfe xnetlib verfügbar.)

Dunford, N, **Schwartz**, J.T.: Linear Operators. Wiley, 1963.

Garabedian, P.R.: Partial Differential Equations. Wiley, 1964.

Golub, G.H., **van Loan**, C.F.: Matrix Computations. North Oxford Academic, Oxford, 1983.

Hackbusch, W.: Multigrid Methods and Applications. Springer, 1985.

Hageman, L.A., **Young**, D.M.: Applied Iterative Methods. Academic Press, 1981.

Jacobi, Carl Gustav (1804-1855): Über eine neue Auflösungsart der bei der Methode der kleinsten Quadrate vorkommenden linearen Gleichungen. Astr.Nachr. 22 Nr. 523 (297-306) 1845.

Michlin, S.G.: Partielle Differentialgleichungen in der mathematischen Physik. Harri Deutsch, 1978.

Minc, H.: Nonnegative Matrices. Wiley, 1988.

Seneta, E.: Nonnegative Matrices, Wiley, 1973.

Smart, D.R.: Fixed Point Theorems. Cambridge, 1974.

Varga, R.S.: Matrix Iterative Analysis. 1962.

Windisch, G.: M-Matrices in Numerical Analysis. Teubner, 1989.

Young, D.: Iterative Solutions of Large Systems, AP 1971.

Kapitel 5

Iterationsverfahren zur Lösung großer linearer Gleichungssysteme: neue Verfahren

Dieses Kapitel befindet sich in Überarbeitung und wird zu einem späteren Zeitpunkt zum Skript hinzugefügt werden.

Kapitel 6

Die Laplace- und Poisson-Gleichung: Methoden der Funktionentheorie

6.1 Zusammenhänge mit der Funktionentheorie

Es besteht ein enger Zusammenhang zwischen der Laplace-Gleichung und der Funktionentheorie.

Definition 6.1 *: (Remmert, S. 45)*
Sei $D \subset \mathbb{C}$ ein Gebiet. Eine Funktion $f : D \to \mathbb{C}$ heißt holomorph in D, wenn f in jedem Punkt von D komplex differenzierbar ist.

Satz 6.1 *: (Remmert, S. 39)*
Notwendig für die komplexe Differenzierbarkeit von $f = u + iv$ in $c \in \mathbb{C}$ ist, daß der Realteil u und der Imaginärteil v von f in c partiell differenzierbar ist und daß die Cauchy- Riemannschen Differentialgleichungen

$$\begin{aligned} \frac{\partial u}{\partial x} &= \frac{\partial v}{\partial y} \\ \frac{\partial u}{\partial y} &= -\frac{\partial v}{\partial x} \end{aligned} \tag{6.1}$$

erfüllt sind.

Es folgt aus diesem Satz, daß - wenn $f = u + iv$ in D holomorph ist und $u, v \in C^2(D)$ - dann gilt

$$\begin{aligned} \Delta u &= u_{xx} + u_{yy} = 0 \\ \Delta v &= v_{xx} + v_{yy} = 0 \, , \end{aligned}$$

d.h. u und v sind Lösungen der Laplace-Gleichungen.

Satz 6.2 : *Maximum-Prinzip für beschränkte Gebiete (Remmert, S. 203)*
Es sei G ein beschränktes Gebiet, und es sei f eine in $\bar{G} := G \cup \partial G$ stetige und in G holomorphe Funktion. Dann nimmt die Funktion $|f|$ ihr Maximum auf dem Rand von G an:

$$\max_{\bar{G}} |f| = \max_{\partial G} |f| .$$

Beweis:
Man stellt fest, daß

$$\begin{aligned} \Delta u^2 &= 2(u_x^2 + u_y^2) \geq 0 \\ \Delta v^2 &= 2(v_x^2 + v_y^2) \geq 0 \end{aligned}$$

und wendet das Maximum-Prinzip (für elliptische Gleichungen) auf die Ungleichung

$$\Delta(u^2 + v^2) \geq 0$$

an.

6.2 Transformation der Laplace-Gleichung (siehe z.B. Kantorowitsch und Krylow, S. 331)

Wir betrachten nochmals die Laplace-Gleichung

$$\Delta_{xy} u = \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0 . \tag{6.2}$$

Es kommt oft vor, daß Probleme für einfache Gebiete wie Kreis, Quadrat, Kreisring, aufgeschnittene Ebene usw. verhältnismäßig einfach gelöst werden können, während sich bei einer komplizierteren Struktur des Gebietes für die unmittelbare Lösung mitunter sehr große Schwierigkeiten ergeben, selbst für ein so einfaches Problem wie das DIRICHLETsche.
Man versucht daher, das gegebene Gebiet vorher auf ein einfacheres abzubilden. Bei derartigen Abbildungen ändern sich im allgemeinen nicht nur das Gebiet, für das die Lösung gesucht ist, und die Randbedingungen, sondern auch die Differentialgleichung, die die gesuchte Funktion erfüllen muß.
Die größte Bedeutung haben für die LAPLACEsche Differentialgleichung offensichtlich die Transformationen, denen gegenüber der LAPLACE- Ausdruck invariant bleibt.
Mit Hilfe der Gleichungen

$$\xi = \xi(x, y) , \quad \eta = \eta(x, y) \tag{6.3}$$

seien für x und y neue Veränderliche ξ und η eingeführt. Wenn die Veränderlichen x, y im Gebiet D variieren, so variieren die neuen Veränderlichen ξ und η in einem gewissen Gebiet G. Wir wollen untersuchen, wie sich die Differentialgleichung bei dieser Transformation ändert:

$$\begin{aligned}
\frac{\partial u}{\partial x} &= \frac{\partial u}{\partial \xi}\frac{\partial \xi}{\partial x} + \frac{\partial u}{\partial \eta}\frac{\partial \eta}{\partial x}\,,\\
\frac{\partial^2 u}{\partial x^2} &= \frac{\partial^2 u}{\partial \xi^2}\left(\frac{\partial \xi}{\partial x}\right)^2 + 2\frac{\partial^2 u}{\partial \xi \partial \eta}\frac{\partial \xi}{\partial x}\frac{\partial \eta}{\partial x} + \frac{\partial^2 u}{\partial \eta^2}\left(\frac{\partial \eta}{\partial x}\right)^2 + \frac{\partial u}{\partial \xi}\frac{\partial^2 \xi}{\partial x^2} + \frac{\partial u}{\partial \eta}\frac{\partial^2 \eta}{\partial x^2}
\end{aligned}$$

und analog

$$\frac{\partial^2 u}{\partial y^2} = \frac{\partial^2 u}{\partial \xi^2}\left(\frac{\partial \xi}{\partial y}\right)^2 + 2\frac{\partial^2 u}{\partial \xi \partial \eta}\frac{\partial \xi}{\partial y}\frac{\partial \eta}{\partial y} + \frac{\partial^2 u}{\partial \eta^2}\left(\frac{\partial \eta}{\partial y}\right)^2 + \frac{\partial u}{\partial \xi}\frac{\partial^2 \xi}{\partial y^2} + \frac{\partial u}{\partial \eta}\frac{\partial^2 \eta}{\partial y^2}\,.$$

Die Differentialgleichung für die Funktion u lautet daher in den Veränderlichen ξ und η

$$\begin{aligned}
\Delta_{xy} u &= \frac{\partial^2 u}{\partial \xi^2}\left[\left(\frac{\partial \xi}{\partial x}\right)^2 + \left(\frac{\partial \xi}{\partial y}\right)^2\right] + 2\frac{\partial^2 u}{\partial \xi \partial \eta}\left(\frac{\partial \xi}{\partial x}\frac{\partial \eta}{\partial x} + \frac{\partial \xi}{\partial y}\frac{\partial \eta}{\partial y}\right) + \\
&+ \frac{\partial^2 u}{\partial \eta^2}\left[\left(\frac{\partial \eta}{\partial x}\right)^2 + \left(\frac{\partial \eta}{\partial y}\right)^2\right] + \frac{\partial u}{\partial \xi}\Delta_{xy}\xi + \frac{\partial u}{\partial \eta}\Delta_{xy}\eta = 0\,.
\end{aligned} \tag{6.4}$$

Damit dies wieder eine LAPLACEsche Differentialgleichung ist, muß offenbar die Transformation (4.3) folgende Forderungen erfüllen:

$$\begin{gathered}
\Delta_{xy}\xi = 0\,; \qquad \Delta_{xy}\eta = 0\,,\\
\frac{\partial \xi}{\partial x}\frac{\partial \eta}{\partial x} + \frac{\partial \xi}{\partial y}\frac{\partial \eta}{\partial y} = 0\,; \qquad \left(\frac{\partial \xi}{\partial x}\right)^2 + \left(\frac{\partial \xi}{\partial y}\right)^2 = \left(\frac{\partial \eta}{\partial x}\right)^2 + \left(\frac{\partial \eta}{\partial y}\right)^2\,.
\end{gathered}$$

Die ersten beiden dieser Differentialgleichungen sagen aus, daß ξ und η harmonische Funktionen von x und y sein müssen. Die dritte Differentialgleichung liefert

$$\frac{\partial \eta}{\partial x} : \frac{\partial \xi}{\partial y} = -\frac{\partial \eta}{\partial y} : \frac{\partial \xi}{\partial x} = -\mu\,; \quad \frac{\partial \eta}{\partial x} = -\mu\frac{\partial \xi}{\partial y}\,; \quad \frac{\partial \eta}{\partial y} = \mu\frac{\partial \xi}{\partial x}\,.$$

Wenn man jetzt in der letzten Differentialgleichung an Stelle der Ableitungen von η diese Ausdrücke einsetzt, so erhält man

$$(1-\mu^2)\left[\left(\frac{\partial \xi}{\partial x}\right)^2 + \left(\frac{\partial \xi}{\partial y}\right)^2\right] = 0\,.$$

Daraus folgt, daß $\mu = \pm 1$ ist, so daß entweder

$$\frac{\partial \xi}{\partial x} = \frac{\partial \eta}{\partial y}\,, \quad \frac{\partial \xi}{\partial y} = -\frac{\partial \eta}{\partial x}\,, \tag{6.5}$$

oder

$$\frac{\partial \xi}{\partial x} = -\frac{\partial \eta}{\partial y}\,, \quad \frac{\partial \xi}{\partial y} = \frac{\partial \eta}{\partial x} \tag{6.6}$$

gilt. Die Differentialgleichungen (4.6) unterscheiden sich von (4.5) lediglich durch Vertauschung von ξ und η und können daher unberücksichtigt bleiben; durch (4.5) wird zum Ausdruck gebracht, daß ξ und η konjugierte harmonische Funktionen sind. Abbildungen, die durch derartige Funktionen vermittelt werden, sind konform, das heißt, im Kleinen ähnlich an allen Stellen der x, y-Ebene, in denen

$$\left(\frac{\partial \xi}{\partial x}\right)^2 + \left(\frac{\partial \xi}{\partial y}\right)^2 \neq 0$$

ist.
Für die Funktion $f(z) = \xi + i\eta$ der komplexen Veränderlichen $z = x + iy$ stellen die Relationen (4.5) die bekannten Regularitätsbedingungen von Cauchy-Riemann dar, und die Laplace-Differentialgleichung erweist sich somit als invariant gegenüber Transformationen, die durch reguläre Funktionen einer komplexen Veränderlichen vermittelt werden. Sei

$$f := \xi + i\eta\,. \tag{6.7}$$

Dann kann die Gleichung (4.4) folgendermaßen geschrieben werden:

$$\Delta_{xy} u = |f'(z)|^2 \Delta_{\xi\eta} u\,. \tag{6.8}$$

Insbesondere, ist $\Delta_{xy} u = 0$, dann gilt auch $\Delta_{\xi\eta} u = 0$.

6.3 Konforme Abbildungen

Definition 6.2 *Sei $f : D \to W$. f heißt konform in D, falls f in D holomorph ist und $f'(z) \neq 0$.*

Bemerkung 6.1 *Sei f eine konforme Abbildung. Sei γ eine Kurve*

$$\gamma : z = z(t)\,, \quad \alpha \leq t \leq \beta\,.$$

Die Abbildung von γ unter f ist

$$\gamma' : w = f(z(t)) \,, \quad \alpha \le t \le \beta \,.$$

Es gilt:

$$w'(t) = f'(z(t)) \cdot z'(t) \,,$$

woraus folgt:

1. $w'(t_0) \neq 0$ *, d.h. γ' hat eine Tangente.*
2. $\arg(w'(t_0)) = \arg(f'(z_0)) + \arg(z'(t_0))$.
3. $\lim_{t\to t_0} \left|\frac{f(z(t))-f(z_0)}{z(t)-z_0}\right| = |f'(z_0)|$.

Zusammenfassend, die Abbildung f ist winkeltreu, und die Länge wird mit dem Faktor $|f'(z)|$ multipliziert.

Sei $u = u(x,y)$, $\Gamma \subset D$ eine Kurve

$$\begin{aligned} u &= \phi_1 \quad \text{auf} \quad \Gamma \\ \frac{\partial u}{\partial n} &= \phi_2 \quad \text{auf} \quad \Gamma \,. \end{aligned}$$

Sei $f : D \to \mathbb{C}$ eine konforme Abbildung $\Gamma' = f(\Gamma)$. Es gilt:

$$\begin{aligned} u &= \phi_1 \quad \text{auf} \quad \Gamma' \\ \frac{\partial u}{\partial n} &= |f'(z)| \cdot \frac{\partial u}{\partial n'} \quad \text{auf} \quad \Gamma' \,. \end{aligned}$$

Der Fundamentalsatz lautet (Kantorowitsch und Krylow, S. 332):
Der einfachste Fall einer konformen Abbildung ist der, daß das abzubildende Gebiet einfach zusammenhängend ist. Die Frage nach der Existenz einer eindeutigen Abbildung zweier solcher Gebiete aufeinander wird durch den Riemannschen Abbildungssatz beantwortet:

Sind D und G einfach zusammenhängende schlichte Gebiete in den Ebenen der komplexen Veränderlichen $z = x + iy$ und $\zeta = \xi + i\eta$, die von der Vollebene und der einfach punktierten Ebene verschieden sind, so existiert eine Funktion $\zeta = f(z)$, die in D holomorph ist und die eine konforme eindeutige Abbildung von D auf G vermittelt. Dabei kann man vorschreiben, daß ein in D gegebener Punkt a in einen willkürlich gewählten Punkt α und eine in a gegebene Richtung in eine bestimmte, dem Punkt α

zugeordnete Richtung übergehen sollen.

Durch die Forderung, daß sich zwei innere Punkte und zwei in diesen Punkten gegebene Richtungen entsprechen sollen, ist die Abbildungsfunktion vollständig bestimmt.
Somit hängt die Funktion, die D auf G abbildet, von drei Parametern ab, da der willkürlich vorgegebene Punkt α durch zwei Parameter und die Richtung in ihm durch einen bestimmt ist. Über die Wahl dieser drei Parameter können wir auch anders verfügen, als im Satz von Riemann angegeben. Beispielsweise kann man für Gebiete, die gewissen Einschränkungen unterworfen sind, welche für die praktisch vorkommenden Fälle sicher erfüllt sind, fordern, daß drei vorgegebene Punkte der Berandung des Gebietes D in drei vorgegebene Randpunkte von G übergehen sollen.

6.4 Die Integralgleichung von Theodorsen

Es gibt mehrere Integralgleichungen, deren Lösung eine konforme Abbildung ermittelt (siehe Gaier, S. 6). Wir geben hier nur eine solche Gleichung an:

Real- und Imaginärteil einer in $|z| < 1$ regulären, in $|z| \leq 1$ stetigen Funktion $f(z)$, genommen auf $|z| = 1$, hängen wie folgt zusammen:

Satz 6.3 *Es sei $f(z) = u(z) + iv(z)(u, v$ reell$)$ in $|z| < 1$ holomorph, in $|z| \leq 1$ stetig. Dann hat $v(e^{i\varphi})$ folgende Integraldarstellung in $u(e^{i\vartheta})$:*

$$v(e^{i\varphi}) = v(0) + \frac{1}{2\pi} \int\limits_{0(H)}^{2\pi} u(e^{i\vartheta}) ctg \, \frac{\varphi - \vartheta}{2} \, d\vartheta \, .$$

Das Integral ist bezüglich $\vartheta = \varphi$ als Cauchyscher Hauptwert zu nehmen.

Beweis: Gaier, S. 62.

Wir formulieren nun unser Problem der konformen Abbildung und zeigen, wie es auf die Lösung einer Integralgleichung zurückgeführt werden kann.
Vorgelegt sei eine bezüglich $w = 0$ sternige Jordankurve C der w-Ebene, und $\varrho = \varrho(\theta)$ sei ihre Darstellung in Polarkoordinaten. Ihr Inneres heiße G, und es sei die konforme Abbildung $w = f(z)$ von $|z| < 1$ auf G zu bestimmen, die durch $f(0) = 0, f'(0) > 0$ normiert ist.
Nach den Ergebnissen von Caratheodory und Osgood ist $f(z)$ in $|z| \leq 1$ stetig und bildet $|z| \leq 1$ eineindeutig auf das abgeschlossene Gebiet $\bar{G}$ ab. Dem Randpunkt $e^{i\varphi}$ entspricht dabei ein Bildpunkt $\varrho(\theta)e^{i\theta}$ mit $\theta = \theta(\varphi)$ auf C. Ist $\theta = \theta(\varphi)$ bekannt, so ist $f(z)$ auf $|z| = 1$ und daher auch für $|z| < 1$ bekannt, etwa auf Grund der Cauchyschen Integralformel. Es ist somit hinreichend, die Ränderzuordnungsfunktion $\theta = \theta(\varphi)$ zu bestimmen.

Diese Funktion $\theta = \theta(\varphi)$ genügt nun einer Integralgleichung. Zu ihrer Ableitung betrachten wir die Hilfsfunktion

$$F(z) = \log \frac{f(z)}{z} = U(z) + iV(z) ;$$

sie ist durch die Vorschrift, daß $F(0) = \log\ f'(0)$ reell sei, eindeutig festgelegt, und sie ist in $|z| < 1$ regulär, in $|z| \le 1$ stetig, ferner $V(0) = 0$.

Weiter ist

$$\begin{array}{ll} U(z) = \log \left|\frac{f(z)}{z}\right| & , \text{ also } U(e^{i\varphi}) = \log\ \varrho(\theta(\varphi)) \\ V(z) = \arg \frac{f(z)}{z} & , \text{ also } V(e^{i\varphi}) = \theta(\varphi) - \varphi , \end{array} \tag{6.9}$$

und Satz 4.6 ergibt daher

$$\theta(\varphi) = \varphi + \frac{1}{2\pi} \int_{0(H)}^{2\pi} \log\ \varrho(\theta(\vartheta)) \text{ctg}\, \frac{\varphi - \vartheta}{2}\, d\vartheta \tag{6.10}$$

oder kurz

$$\theta(\varphi) = \varphi + K[\log\ \varrho(\theta(\varphi))] .$$

Dies ist die nichtlineare, singuläre Integralgleichung von Theodorsen zur Ermittlung der gesuchten Ränderzuordnungsfunktion $\theta = \theta(\varphi)$.

6.5 Das Schwarzsche Spiegelungsprinzip

Satz 6.4 *: (siehe z.B. Goluzin, S. 44)*
Eine eindeutige konforme Abbildung f eines Gebiets D mit Jordanrand ∂D auf dem offenen Einheitskreis G läßt sich als eine stetige eindeutige Abbildung von $\bar{D}$ auf $\bar{G}$ erweitern.

Ist $\gamma \subset \partial D$ eine offene analytische Jordankurve und $z_0 \in \gamma$, dann ist f holomorph in z_0 und $f'(z_0) \neq 0$.

Satz 6.5 *: (Spiegelungsprinzip von Schwarz)*
Seien D_z und D_w zwei Gebiete mit $\alpha_z \in \partial D_z$ und $\alpha_w \in \partial D_w$, wo α_z und α_w Geradensegmente sind. Sei $w = f(z)$ eine konforme Abbildung von D_z auf D_w. Sei $f(\alpha_z) = \alpha_w$. Dann läßt sich f über α_z hinaus wie folgt analytisch fortsetzen:

Spiegelt man D_z an α_z und ordnet man dem Spiegelbild z^* von z den durch Spiegelung von $w = F(z)$ und α_w hervorgegangenen Wert w zu, so ist die so definierte

Funktion $w^* = f(z^*)$ in dem gespiegelten Gebiet D_w^* analytisch und somit eine analytische Fortsetzung von f.

Beweis: (siehe z.B. Nehari, S. 184)

O.E. nehmen wir an, daß α_z und α_w Intervalle der reellen Achse sind. D.h., $z^* = \bar{z}$, $w^* = \bar{w}$.

Die Funktion

$$g(z) := \overline{f(\bar{z})}$$

ist holomorph in D_z^* wie man z.B. aus der Cauchy-Riemann- Gleichung sieht. Auf α_z gilt:

$$\begin{aligned} \bar{z} &= z \, , \\ \overline{f(\bar{z})} &= f(\bar{z}) = f(z) \, , \end{aligned}$$

da $f(z)$ dort reell ist.

Durch Addition der Integrale

$$\int_{\partial D_z} \frac{f(\zeta)}{\zeta - z} \, d\zeta$$

und

$$\int_{\partial D_z^*} \frac{g(\zeta)}{\zeta - z} \, d\zeta$$

folgt, daß die Funktion

$$F(z) = \begin{cases} f(z) & , \quad z \in D_z \cup \alpha_z \\ g(z) & , \quad z \in D_z^* \cup \alpha_z \end{cases}$$

in $D_z^* \cup \alpha_z \cup D_z$ holomorph ist.

6.6 Die Formel von Christoffel-Schwarz

6.6.1 Herleitung (siehe Kantorowitsch und Krylow, S. 475, und Nehari, S. 189)

Es sei $w = f(z)$ eine Funktion, die eine konforme Abbildung der oberen z-Halbebene D auf das Innere G eines gegebenen Polygons A der w-Ebene mit den Innenwinkeln $\alpha_1\pi, \alpha_2\pi, \ldots, \alpha_n\pi$ und den Eckpunkten $A_1, A_2, \ldots, A_n$ vermittelt. Die Punkte $A_{nu}(\nu = 1, \ldots, n)$ entsprechen dann gewissen Punkten $a_1, a_2, \ldots, a_n$ der reellen Achse der z-Ebene, von denen wir annehmen wollen, daß sie alle im Endlichen liegen und in der Ordnung

$$a_1 < a_2 < \cdots < a_n$$

aufeinanderfolgen
Der Rand ∂G des Polynoms ist eine Jordan-Kurve, so daß sich (Satz 4.6 im vorherigen Absatz) die Abbildung f stetig und eindeutig erweitern läßt:

$$f : D \cup \partial D \to G \cup \partial G \ .$$

Das Intervall $\alpha = (a_k, a_{k+1})$ ist eine analytische Jordan-Kurve, so daß f auf α holomorph und konform ($f' \neq 0$) ist.

Sei nun $\alpha_z^{(1)} := (a_1, a_2)$, $\alpha_w^{(1)} = f(\alpha_z^{(1)})$. Wegen des Schwarzschen Spiegelungsprinzips läßt sich f durch Spiegelung auf der unteren Halbebene D^* analytisch fortsetzen. Sei f_1 die dadurch gewonnene Funktion:

$$f_1 : D^* \to G^* \ .$$

Sei $\alpha_z^{(2)} := (a_2, a_3)$, $\alpha_w^{(2)} = f_1(\alpha_z^{(2)})$.
Wegen des Schwarzschen Spiegelungsprinzips läßt sich f_1 durch Spiegelung auf der oberen Halbebene $(D^*)^* = D$ analytisch fortsetzen. Sei f_2 diese Funktion:

$$f_2 : D \to (G^*)^* \ .$$

Das Polygon G kann durch eine Drehung und eine Verschiebung auf $(G^*)^*$ abgebildet werden:

$$\begin{aligned} f_3 \quad : \quad & G \to (G^*)^* \ , \\ & z \to Az + B \ , \end{aligned}$$

wo A und B komplexe Konstanten sind.

Es folgt:

$$f_2(z) = (f_3 \circ f)(z) = Af(z) + B \ .$$

Sei g die 'Schwarzsche Funktion'

$$g(z) := \frac{f''(z)}{f'(z)} .$$

Sei

$$H := \mathbb{C} \backslash \{a_1, \ldots, a_n\} .$$

Da

$$\frac{f_z''(z)}{f_z'(z)} = \frac{f''(z)}{f'(z)}$$

folgt, daß durch die o.a. analytische Fortsetzung von f, g als eindeutige holomorphe Funktion in H definiert wird.

Wir werden jetzt das Verhalten der Funktion $g(z)$ in der Umgebung der Punkte $a_1, \ldots, a_n$ untersuchen. Sei

$$h(z) := [f(z) - f(a_1)]^{\frac{1}{\alpha_1}} .$$

Sei

$$\begin{aligned} U_\epsilon &:= \{z = x + iy : y > 0 \text{ und } |z - a_1| < \epsilon\} , \\ \alpha_z^{(1)} &:= \{x : -\epsilon < x - \mathbb{R}a_1 \leq 0\} , \\ \alpha_z^{(2)} &:= \{x : 0 \leq x - \mathbb{R}a_1 < \epsilon\} . \end{aligned}$$

Es gilt:

$$\begin{aligned} f(\alpha_z^{(1)}) &= A_n A_1 , \\ f(\alpha_z^{(2)} &= A_1 A_2 , \end{aligned}$$

so daß h das Geradensegment $\alpha_z^{(1)} \cup \alpha_z^{(2)}$ auf einem Geradensegment abbildet. Aufgrund des Schwarzschen Spiegelungsprinzips folgt dann, daß h eine volle Umgebung des Punktes a_1 konform und eineindeutig auf eine volle Umgebung des Punktes 0 abbildet. D.h. h ist in $z = a_1$ holomorph.

Sei nun $\mu_1 := 1 - \alpha_1$. Es folgt

$$f(z) = f(a_1) + h(z)^{1-\mu_1}$$

mit $h(z)$ holomorph in $z = a_1$ und $h'(a_1) \neq 0$. D.h.

$$h(z) = (z - a_1)h_1(z)$$

mit h_1 holomorph und $h_1(a_1) \neq 0$. Es folgt

$$\frac{f''(z)}{f'(z)} = \frac{-\mu_1}{(z - a_1)} + k(z)$$

mit $k(z)$ holomorph in $z = a_1$. Nach wiederholter Benutzung dieses Arguments können wir folgern, daß

$$g_1(z) := g(z) + \sum_{k=1}^{n} \frac{\mu_k}{k - a_k}$$

in $\mathbb{C}$ holomorph ist. Nach dem Satz von Liouville ist $g_1(z)$ konstant.

In einer Umgebung von $z = \infty$ gibt es eine Entwicklung von $f(z)$,

$$f(z) = f(\infty) + \frac{c_1}{z} + \frac{c_2}{z^2} + \cdots ,$$

wodurch leicht einzusehen ist, daß

$$\frac{f''(z)}{f'(z)} = O\left(\frac{1}{z}\right) .$$

Wir können sofort daraus schließen, daß $g(\infty) = g_1(\infty) = 0$, so daß

$$g(z) = \frac{f''(z)}{f'(z)} = \sum_{k=1}^{n} \frac{1 - \alpha_k}{z - a_k} .$$

Durch zweimalige Integration erhält man nacheinander

$$\begin{aligned} ln\ f'(z) &= \sum_{k=1}^{n} (\alpha_k - 1) ln(z - a_k) + ln\ C_1 = ln\ C_1 (z - a_1)^{\alpha_1 - 1} \cdots (z - a_n)^{\alpha_n - 1} , \\ f(z) &= C_1 \int_0^z (v - a_1)^{\alpha_1 - 1} (v - a_2)^{\alpha_2 - 1} \cdots (v - a_n)^{\alpha_n - 1} dv + C_2 . \end{aligned}$$

Das ist die gesuchte Darstellung der Abbildung $f(z)$. Diese Formel bezeichnet man gewöhnlich als die Formel von Christoffel-Schwarz.

6.6.2 Die Parameterwerte des Christoffel-Schwarzschen Integrals (siehe Kantorowitsch und Krylow, S. 477)

Die Funktion

$$w = C_1 \int_0^z (v-a_1)^{\alpha_1-1}(v-a_2)^{\alpha_2-1}\cdots(v-a_n)^{\alpha_n-1}dv + C_2 \;, \tag{6.11}$$

welche die obere z-Halbebene auf das Innere eines Polygons der w-Ebene abbildet, enthält $2n+2$ Parameter $\alpha_1, \alpha_2, \dots, \alpha_n$, $a_1, \dots, a_n, C_1, C_2$. Sind die Werte aller dieser Parameter bekannt, so ist auch das Polygon bestimmt. Die Innenwinkel des Polygons sind gleich $\alpha_1, \pi, \alpha_2\pi, \dots, a_n\pi$. Die Koordinaten b_k der Eckpunkte A_k erhält man aus den Gleichungen

$$b_k = C_1 \int_0^{a_k} (z-a_1)^{\alpha_1-1}(z-a_2)^{\alpha_2-1}\cdots(z-a_n)^{\alpha_n-1}dt + C_2 \;, \; 1 \le k \le n \;. \tag{6.12}$$

Für die Länge der Seite A_kA_{k+1} berechnet man

$$\begin{aligned} A_kA_{k+1} = |b_{k+1}-b_k| = \Bigg| C_1 \int_0^{a_{k+1}} (z-a_1)^{\alpha_1-1}(z-a_2)^{\alpha_2-1}\cdots(z-a_n)^{\alpha_n-1}dt + \\ +C_2 - \left[C_1 \int_0^{a_k} (z-a_1)^{\alpha_1-1}(z-a_2)^{\alpha_2-1}\cdots(z-a_n)^{\alpha_n-1}dt + C_2 \right] \Bigg| = \\ = |C_1| \int_{a_k}^{a_{k+1}} (z-a_1)^{\alpha_1-1}(z-a_2)^{\alpha_2-1}\cdots(z-a_k)^{\alpha_k-1}(a_{k+1}-z)^{\alpha_{k+1}-1}\cdots \times \\ \times \cdots (a_n-z)^{\alpha_n-1}dt \quad . \end{aligned} \tag{6.13}$$

In der Praxis ist gewöhnlich eine Funktion gesucht, die eine konforme Abbildung einer Halbebene auf das Innere eines vollständig bestimmten Polygons vermittelt. Die Parameter

$$a_1, a_2, \dots, a_n \;, \quad C_1, C_2$$

in der Abbildungsformel müssen dann in entsprechender Weise gewählt werden. Das Verhältnis der 2-ten, 3-ten,…, $(n-2)$-ten Seite zur ersten Seite bezeichnen wir mit $\lambda_2, \lambda_3, \dots, \lambda_{n-2}$.

$$\frac{\text{Seitenlänge } A_kA_{k+1}}{\text{Seitenlänge } A_1A_2} = \lambda_k \;, \quad 2 \le k \le n-1 \;. \tag{6.14}$$

Wählen wir in der Funktion

$$w^* = \int_0^z (v-a_1)^{\alpha_1-1}(v-a_2)^{\alpha_2-1}\cdots(v-a_n)^{\alpha_n-1}dv$$

die $a_1, a_2, \ldots, a_n$ so, daß die Relationen (4.14) erfüllt sind, so vermittelt diese Funktion eine konforme Abbildung der oberen Halbebene auf das Innere eines zu G ähnlichen Polygons G^*. Durch eine lineare Abbildung der Form

$$w = C_1 w^* + C_2$$

kann das Polygon G^* auf G jetzt abgebildet werden.
Zur Bestimmung der Parameter $a_1, a_2, \ldots, a_n$ haben wir nur $n-3$ Gleichungen zur Verfügung. Daher können drei Parameter völlig willkürlich gewählt werden. Dies entspricht der Tatsache, daß wir in dem Integral (4.11) eine linear gebrochene Substitution vornehmen können, welche die obere Halbebene so in sich transformiert, daß drei Punkte der reellen Achse in drei beliebig vorgegebene Punkte derselben übergehen. Diese willkürlichen Punkte seien $z = a_1,\ z = a_2$ und $z = a_n$. Wir bezeichnen sie im Unterschied zu den anderen Punkten, die bestimmt werden müssen, mit p_1, p_2 und p_3. Die Relationen (4.14) lauten ausführlich geschrieben:

$$\left.\begin{array}{lcl} I_2(a_3, a_4, \ldots, a_{n-1}) & = & \lambda_2 I_1(a_3, a_4, \ldots, a_{n-1}), \\ I_3(a_3, a_4, \ldots, a_{n-1}) & = & \lambda_3 I_1(a_3, a_4, \ldots, a_{n-1}), \\ \multicolumn{3}{c}{\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots} \\ I_{n-2}(a_3, a_4, \ldots, a_{n-1}) & = & \lambda_{n-2} I_1(a_3, a_4, \ldots, a_{n-1}) \end{array}\right\} \qquad (6.15)$$

mit

$$\left.\begin{array}{lcl} I_1 & = & \int\limits_{p_1}^{p_2} (z-p_1)^{\alpha_1-1}(p_2-z)^{\alpha_2-1}(a_3-z)^{\alpha_3-1}\ldots\times \\ & & \times\ldots(a_{n-1}-z)^{\alpha_{n-1}-1}(p_3-z)^{\alpha_n-1}dt, \\ I_2 & = & \int\limits_{p_2}^{a_3} (z-p_1)^{\alpha_1-1}(z-p_2)^{\alpha_2-1}(a_3-z)^{\alpha_3-1}\ldots\times \\ & & \times\ldots(a_{n-1}-z)^{\alpha_{n-1}-1}(p_3-z)^{\alpha_n-1}dt, \\ \multicolumn{3}{c}{\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots} \\ I_{n-2} & = & \int\limits_{a_{n-2}}^{a_{n-1}} (z-p_1)^{\alpha_1-1}(z-p_2)^{\alpha_2-1}(z-a_3)^{\alpha_3-1}\ldots\times \\ & & \times\ldots(z-a_{n-2})^{\alpha_{n-2}-1}(a_{n-1}-z)^{\alpha_{n-1}-1}(p_3-z)^{\alpha_n-1}dt. \end{array}\right\} \qquad (6.16)$$

6.6.3 Die Christoffel-Schwarz-Formel: Ein Beispiel

Das Polygon G sei ein Dreieck mit den Winkeln $\alpha_1\pi, \alpha_2\pi, \alpha_3\pi$. Da zwei Dreiecke mit gleichen Winkeln immer ähnlich sind, können wir in diesem Falle die Konstanten a_k beliebig wählen. Wir setzen $a_1 = 0,\ a_2 = 1,\ a_3 = \infty$. Die Abbildungsfunktion hat die Form

$$w = C_1 \int\limits_0^z v^{\alpha_1-1}(1-v)^{\alpha_2-1}dv + C_2\ .$$

Zu ihrer vollständigen Bestimmung ist es notwendig, die Konstanten C_1 und C_2 zu ermitteln. Wir bestimmen C_1 und C_2 für ein gleichschenklig-rechtwinkliges Dreieck, dessen Ecken die Punkte $A_1(0), A_2(1)$ und $A_3(i)$ sind.
Dem Punkt $z = 0$ möge der Eckpunkt A_1 und dem Punkt $z = 1$ der Eckpunkt A zugeordnet sein. Dann ist $\alpha_1 = \frac{1}{2}$, $\alpha_2 = \frac{1}{4}$, und die Abbildungsfunktion lautet

$$w = C_1 \int\limits_0^z v^{-\frac{1}{2}}(1-v)^{-\frac{3}{4}} dv + C_2 \ .$$

Da nach Voraussetzung $z = 0$ in $w = 0$ übergeht, ist $C_2 = 0$. Für $z = 1$ müssen wir die Koordinate des Eckpunktes A_2 erhalten. Somit ergibt sich für C_1 die Gleichung

$$C_1 \int\limits_0^1 t^{-\frac{1}{2}}(1-z)^{-\frac{3}{4}} dt = 1 \ ,$$

das heißt, es ist

$$C_1 = \frac{1}{\int_0^1 z^{-\frac{1}{2}}(1-z)^{-\frac{3}{4}} dt} \ .$$

Wir wollen das im Nenner stehende Integral

$$I = \int\limits_0^1 t^{-\frac{1}{2}}(1-z)^{-\frac{3}{4}} dt = \int\limits_0^{\frac{1}{2}} z^{-\frac{1}{2}}(1-z)^{-\frac{3}{4}} dt + \int\limits_{\frac{1}{2}}^1 z^{-\frac{1}{2}}(1-z)^{-\frac{3}{4}} dt = I_1 + I_2$$

berechnen.
Zunächst betrachten wir I_1. Den Integranden stellen wir in der Form

$$z^{-\frac{1}{2}}(1-z)^{-\frac{3}{4}} = z^{-\frac{1}{2}}\left(1 + \frac{3}{4}\,z\right) + z^{-\frac{1}{2}}\left[(1-z)^{-\frac{3}{4}} - 1 - \frac{3}{4}\,z\right]$$

dar und erhalten

$$I_1 = \int\limits_0^{\frac{1}{2}} z^{-\frac{1}{2}}\left(1 + \frac{3}{4}\,z\right) dt + \int\limits_0^{\frac{1}{2}} z^{-\frac{1}{2}}\left[(1-z)^{-\frac{3}{4}} - 1 - \frac{3}{4}\,z\right] dt \ .$$

Das erste dieser Integrale läßt sich explizit berechnen; sein Wert ist gleich $1,5910$. Das zweite berechnen wir nach der Simpsonschen Formel. Wir bezeichnen

$$z^{-\frac{1}{2}}\left[(1-z)^{-\frac{3}{4}} - 1 - \frac{3}{4}\,z\right]$$

mit $y(z)$; dann ist

$$\begin{array}{lll} y(0) = 0, & y\left(\frac{1}{16}\right) = 0,0104, & y\left(\frac{1}{8}\right) = 0,0328, \\ y\left(\frac{3}{16}\right) = 0,0644, & y\left(\frac{1}{4}\right) = 0,1066, & y\left(\frac{5}{16}\right) = 0,1612, \\ y\left(\frac{3}{8}\right) = 0,2309, & y\left(\frac{7}{16}\right) = 0,3198, & y\left(\frac{1}{2}\right) = 0,4339, \\ & \int\limits_0^{\frac{1}{2}} y(z)dt = 0,0707, & \\ & I_1 = 1,5910 + 0,0707 = 1,6617. & \end{array}$$

Nun gehen wir zur Berechnung von I_2 über. Den Integranden stellen wir in der Form

$$z^{-\frac{1}{2}}(1-z)^{-\frac{3}{4}} = (1-z)^{-\frac{3}{4}}\left[1+\frac{1}{2}(1-z)\right] + (1-z)^{-\frac{3}{4}}\left[z^{-\frac{1}{2}} - 1 - \frac{1}{2}(1-z)\right]$$

dar und erhalten

$$I_2 = \int\limits_{\frac{1}{2}}^{1}(1-z)^{-\frac{3}{4}}\left[1+\frac{1}{2}(1-z)\right]dt + \int\limits_{\frac{1}{2}}^{1}(1-z)^{-\frac{3}{4}}\left[z^{-\frac{1}{2}} - 1 - \frac{1}{2}(1-z)\right]dt\,.$$

Der Wert des ersten Integrals ist gleich 3,5318.
Das zweite Integral läßt sich ebenso berechnen wie der zweite Teil des Integrals I_1. Man findet hierfür den Wert 0,0506, so daß

$$I_2 = 3,5824, \quad I = I_1 + I_2 = 5,2441$$

und somit

$$C_1 = \frac{1}{5,2441} = 0,1907$$

ist. Das Problem ist damit vollständig gelöst.

Um unser Resultat nachzuprüfen, berechnen wir noch die Hypotenuse des Dreiecks

$$\begin{aligned} A_2A_3 &= 0,1907\int\limits_1^{\infty} z^{-\frac{1}{2}}(z-1)^{-\frac{3}{4}}dt\,; \\ \int\limits_1^{\infty} z^{-\frac{1}{2}}(z-1)^{-\frac{3}{4}}dt &= 2\int\limits_0^{\frac{1}{2}} \tau^{-\frac{3}{4}}(1-\tau)^{-\frac{3}{4}}d\tau \qquad \left(\tau = \frac{1}{z}\right)\,; \\ \int\limits_0^{\frac{1}{2}} \tau^{-\frac{3}{4}}(1-\tau)^{-\frac{3}{4}}d\tau &= \int\limits_0^{\frac{1}{2}} \tau^{-\frac{3}{4}}\left(1+\frac{3}{4}\tau\right)d\tau + \int\limits_0^{\frac{1}{2}} \tau^{-\frac{3}{4}}\left[(1-\tau)^{-\frac{3}{4}} - 1 - \frac{3}{4}\tau\right]d\tau = \\ &= 3,6158 + 0,0922 = 3,7080\,;\, A_2A_3 = 0,1907\cdot 2\cdot 3,7080 = 1,4142\,. \end{aligned}$$

Der genaue Wert von A_2A_3 ist gleich $\sqrt{2} = 1,4142\ldots.$

Literatur

Gaier, D.: Konstruktive Methoden der konformen Abbildung. Berlin: Springer, 1964.

Kantorowitsch, L.W., **Krylow**, W.I.: Näherungsmethoden der höheren Analysis. Berlin: VEB Deutscher Verlag der Wissenschaften, 1956.

Nehari, Z.: Conformal Mapping. New York.

Remmert, R.: Funktionentheorie I (zweite Ausgabe). Berlin: Springer, 1989.

Schwarz, H.A.: Über einige Abbildungsaufgaben. Ges.Math. Abh. 1869 Bd. II, Seite 65-83.

Trefethen, L.N. (Herausgeber): Numerical Conformal Mapping. Amsterdam, North-Holland, 1986.

Kapitel 7

Die Laplace-Gleichung: Integralgleichungen

7.1 Einleitung

Die Lösung partieller Differentialgleichungen mit Hilfe von Integralgleichungen hat eine sehr lange Geschichte. In der Tat sind solche Methoden schon vor der Herleitung der Theorie partieller Differentialgleichungen benutzt worden. Der erste Beweis, daß das Dirichlet-Problem für die Laplace-Gleichung lösbar ist, wurde durch Neumann 1866(?) gegeben, und er benutzt Integralgleichungen. Siehe Enzyklopädie der Mathematik wegen der früheren Geschichte dieses Gebietes.
Die numerische Lösung von Randwertaufgaben durch die numerische Lösung von Integralgleichungen ist auch sehr lange bekannt. Z.B. hat Trefftz (1913) die Strömung eines Wasserstrahls auf diese Weise gelöst. In letzter Zeit ist dieses Verfahren als die "Boundary-Element" Methode bekannt geworden.

7.2 Integralgleichungen für die Laplace-Gleichung

Wir betrachten das Dirichlet-Randwertproblem in zwei Dimensionen:

$$\begin{aligned} \Delta u &= u_{xx} + u_{yy} = 0 \,, \quad \text{in } \Omega \,, && (7.1) \\ u &= f \quad \text{auf } \Gamma = \partial\Omega \,. && (7.2) \end{aligned}$$

Es gibt drei bekannte Darstellungen der Lösung u von (5.1) als Integrale auf $\Gamma := \partial\Omega$.

7.2.1 Die Integralgleichung für eine Doppelschicht

(Siehe Kantorowitsch und Krylow, S. 115.)
Bezeichnet man mit r den Abstand zwischen den Punkten $z = x + iy = P(x, y)$ und $\zeta = \xi + i\eta = M(\xi, \eta)$, so ist

$$
\begin{aligned}
u := \Re(ln(z-\zeta)) = ln\ r &= ln\ \sqrt{(x-\xi)^2+(y-\eta)^2} \\
&= \frac{1}{2}\, ln[(x-\xi)^2+(y-\eta)^2]
\end{aligned}
\tag{7.3}
$$

eine Lösung der Laplace-Differentialgleichung

$$
\Delta u = \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0\ . \tag{7.4}
$$

Die Ableitungen der Funktion u nach den Parametern ξ und η sind ebenfalls harmonisch, beispielsweise gilt für $\frac{\partial u}{\partial \xi}$

$$
\frac{\partial^2}{\partial x^2}\left(\frac{\partial u}{\partial \xi}\right) + \frac{\partial^2}{\partial y^2}\left(\frac{\partial u}{\partial \xi}\right) = \frac{\partial}{\partial \xi}\left(\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2}\right) = 0\ .
$$

Daraus folgt, daß auch die Ableitung von u (als Funktion der Variablen ξ und η) in einer beliebigen konstanten Richtung L, die eine Linearkombination von $\frac{\partial u}{\partial \xi}$ und $\frac{\partial u}{\partial \eta}$ ist, ebenfalls eine harmonische Funktion darstellt.

Wenn man mit φ den Winkel bezeichnet, den der Vektor $\vec{PM}$ mit der Richtung L bildet, so lautet diese Ableitung

$$
\frac{\partial u}{\partial L} = \frac{1}{r}\frac{\partial r}{\partial L} = \frac{\cos\ \varphi}{r}\ .
$$

Also ist $\frac{\cos\ \varphi}{r}$ eine harmonische Funktion.

Nach diesen einleitenden Bemerkungen kommen wir zur Untersuchung des Dirichlet-Problems. Es sei ein endliches Gebiet Ω gegeben, das durch eine einfach geschlossene Kurve $\partial\Omega$ begrenzt ist. Wir nehmen an, daß die Berandung in der Parameterdarstellung

$$
x = x(s)\ ,\ y = y(s) \quad (0 \le s \le s_0)
$$

gegeben ist; als Parameter verwenden wir die Bogenlänge s. Wir setzen voraus, daß die Funktionen $x(s)$ und $y(s)$ stetige Ableitungen $x'(s)$ und $y'(s)$ besitzen, die nicht gleichzeitig Null werden. Es soll eine Funktion u gefunden werden, die in Ω harmonisch ist und auf der Berandung in eine vorgegebene stetige Funktion $f(s)$ der Bogenlänge s übergeht, das heißt, es sei

$$
u = f(s) \quad \text{auf}\ \partial\Omega\ .
$$

Wir wählen einen beliebigen Punkt $M(\xi,\eta)$ auf $\partial\Omega$ und bezeichnen mit φ den Winkel, der durch den Vektor $\vec{PM}$ und die äußere Normale in M gebildet wird. Nach dem oben

Gesagten ist $\frac{\cos\varphi}{r}$ eine harmonische Funktion der Veränderlichen x und y in Ω. Es sei jetzt $\mu(s)$ eine willkürliche stetige Funktion. Dann ist auch das Integral

$$V(x,y) := \int_0^{s_0} \frac{\cos\,\varphi}{r}\,\mu(s)ds \tag{7.5}$$

eine im Gebiet Ω harmonische Funktion der Veränderlichen x und y.

Die Funktion V nennt man das Potential einer Doppelschicht, und $\mu(s)$ heißt seine Dichte. Wir geben V eine andere Form, um die geometrische Bedeutung klarer hervortreten zu lassen. Mit ω bezeichnen wir den Winkel, den der Vektor $\vec{PM}$ mit der positiven x-Achse bildet. Wenn man die Lage des Punktes M ändert, so ist ω eine Funktion von s, und $d\omega$ ist der Winkel, unter dem das Bogenelement ds von P aus gesehen wird. Die Projektion von ds auf das Lot zum Vektor $\vec{PM}$ ist $ds\ \cos\,\varphi$. Daher ist der Winkel, unter dem ds vom Punkte P aus erscheint, gleich $\frac{\cos\,\varphi\, ds}{r}$, also ist $\frac{\cos\,\varphi}{r}\,ds = d\omega$. Man kann somit V die Gestalt

$$V = \int_0^{s_0} \frac{\cos\,\varphi}{r}\,\mu(s)ds = \int_{\partial\Omega} \mu(s)d\omega = \int_0^{s_0} \frac{d\omega}{ds}\,\mu(s)ds \tag{7.6}$$

geben. Dabei ist ω eine Funktion von s, x und y, also $\omega = \omega(s,x,y)$; s_0 bezeichnet die Länge $\partial\Omega$.
Haben wir insbesondere die Dichte $\mu(s) = 1$, dann liefert $V = \int_0^{s_0} d\omega$ offensichtlich den Winkel, unter dem die gesamte Kurve $\partial\Omega$ vom Punkte P aus gesehen wird. Es ist klar, daß dieser Winkel gleich 2π ist, wenn P innerhalb $\partial\Omega$ liegt; er ist gleich π, wenn P ein Punkt der Berandung $\partial\Omega$ ist, und schließlich gleich Null, wenn P außerhalb $\partial\Omega$ liegt:

$$\int_0^{s_0} \frac{d\omega}{ds}\,ds = \begin{cases} 2\pi & ,\ \text{falls } P \text{ innerhalb } \partial\Omega\ , \\ \pi & ,\ \text{falls } P \text{ auf } \partial\Omega\ , \\ 0 & ,\ \text{falls } P \text{ außerhalb } \partial\Omega\ . \end{cases} \tag{7.7}$$

Wir versuchen jetzt, eine Lösung des Dirichtlet-Problems in Form eines Potentials einer Doppelschicht zu ermitteln. Dazu machen wir den Ansatz

$$u(x,y) = \int_0^{s_0} \mu(s)\,\frac{d\omega}{ds}\,ds \tag{7.8}$$

und suchen die unbekannte Dichte $\mu(s)$ derart zu bestimmen, daß die Funktion $u(x,y)$ bei Annäherung an die Berandung gegen die Werte von $f(s)$ konvergiert.

Wir wollen also feststellen, welchen Grenzwert $u(x,y)$ hat, wenn $P(x,y)$ gegen einen Punkt $P_0(x(\sigma), y(\sigma)) \in \partial\Omega$ strebt. Man erhält wegen $\int\limits_0^{s_0} \mu(\sigma) d\omega(s,\sigma) = \pi\mu(\sigma)$,

$$\begin{aligned}
\lim_{P\to P_0} u(x,y) &= \lim_{P\to P_0} \int\limits_0^{s_0} \mu(s) d\omega \\
&= \lim_{P\to P_0} \left[\int\limits_0^{s_0} [\mu(s) - \mu(\sigma)] \tfrac{d\omega(s,x,y)}{ds}\, ds + \mu(\sigma) \int\limits_0^{s_0} \tfrac{d\omega(s,x,y)}{ds}\, ds \right] \\
&= \lim_{P\to P_0} \left[\int\limits_0^{\sigma-\epsilon} [\mu(s) - \mu(\sigma)] \tfrac{d\omega(s,x,y)}{ds}\, ds \right. \\
&\quad + \int\limits_{\sigma-\epsilon}^{\sigma+\epsilon} [\mu(s) - \mu(\sigma)] \tfrac{d\omega(s,x,y)}{ds}\, ds \\
&\quad + \int\limits_{\sigma+\epsilon}^{s_0} [\mu(s) - \mu(\sigma)] \tfrac{d\omega(s,x,y)}{ds}\, ds \\
&\quad \left. + \mu(\sigma) \int\limits_0^{s_0} \tfrac{d\omega(s,x,y)}{ds}\, ds \right], \\
&= \int\limits_0^{s_0} [\mu(s) - \mu(\sigma)] \tfrac{d\omega(s,\sigma)}{ds}\, ds + 2\pi\mu(\sigma) \\
&= \int\limits_0^{s_0} \mu(s) \tfrac{d\omega(s,\sigma)}{ds}\, ds + \pi\mu(\sigma)
\end{aligned} \tag{7.9}$$

mit

$$\omega(s,\sigma) := \omega(s, x(\sigma), y(\sigma)) \ .$$

Also muß

$$\pi\mu(\sigma) + \int\limits_0^{s_0} \mu(s) \frac{d\omega}{ds}\, ds = f(\sigma) \tag{7.10}$$

gelten.

Die erhaltene Relation ist eine Integralgleichung für die unbekannte Funktion $\mu(s)$. Wenn man durch π dividiert, so bekommt man

$$\mu(\sigma) - \int\limits_0^{s_0} K(s,\sigma)\mu(s) ds = \frac{1}{\pi}\, f(\sigma) \ . \tag{7.11}$$

Die Gleichung (5.11) ist eine Fredholmsche Integralgleichung zweiter Art.

Der Kern $K(s,\sigma)$ kann, wenn man den geometrisch evidenten Ausdruck für den Winkel $\omega(s,\sigma)$ benutzt, in der Form

$$K(s,\sigma) = -\frac{1}{\pi}\, \frac{d\omega}{ds} = -\frac{1}{\pi}\, \frac{d}{ds} \operatorname{arctg} \frac{y(s) - y(\sigma)}{x(s) - x(\sigma)} \tag{7.12}$$

geschrieben werden.

Wenn man in (5.11) vor dem Integralzeichen einen Parameter λ einführt,

$$\mu(\sigma) - \lambda \int_0^{s_0} K(s,\sigma)\mu(s)ds = \frac{1}{\pi} f(\sigma) , \tag{7.13}$$

so kann man zeigen, daß der Wert $\lambda = 1$, der für uns interessant ist, kein Eigenwert für die gegebene Integralgleichung ist. Daher hat (5.11) auf Grund der allgemeinen Theorie der Integralgleichungen für beliebiges $f(\sigma)$ genau eine Lösung.
Ist $\partial\Omega$ konvex, so folgt die Behauptung, daß $\lambda = 1$ kein Eigenwert ist und zwar aus der Tatsache, daß

$$K(s,\sigma) \leq 0 , \quad \int_0^{s_0} K(s,\sigma)ds = 1 ,$$

wie schon von Neumann bemerkt. Der Beweis im allgemeinen Fall wurde erst von Fredholm im Jahre 1906 erbracht.
Für die angenäherte Bestimmung der Lösung spielt die Verteilung der Eigenwerte dieser Gleichung eine wesentliche Rolle. Vor allem kann man zeigen, daß diese reell sind. Ferner ist $\lambda = -1$ ein Eigenwert, da eine Konstante $\mu(s) = C$ offenbar die zu (5.13) mit $\lambda = -1$ gehörige homogene Integralgleichung erfüllt. Es gilt nämlich unter Benutzung von (5.7)

$$\pi C - \int_0^{s_0} C \frac{d\omega(s,\sigma)}{ds} \, ds = 0 .$$

Schließlich kann man feststellen, daß es im Intervall $(-1, 1)$ keine Eigenwerte gibt.

Zuletzt schreiben wir (5.11) in noch einer Form. Sei $\mathbf{x} = (x.y)$, $\mathbf{y} = (y_1, y_2)$, $\mathbf{n_y} =$ äußere Normale zu Γ in der Stelle $\mathbf{y} \in \Gamma$, $\mathbf{r} = \mathbf{y} - \mathbf{x}$.

Es läßt sich feststellen, daß

$$\begin{aligned} & \frac{\partial}{\partial \mathbf{n_y}} \ln |\mathbf{x} - \mathbf{y}| \\ = \; & \left(n_1 \frac{\partial}{\partial y_1} + n_2 \frac{\partial}{\partial y_2} \right) \ln |\mathbf{x} - \mathbf{y}| , \\ = \; & (n_1, n_2) \cdot \mathbf{r}/|\mathbf{r}|^2 , \\ = \; & \cos \varphi/|\mathbf{r}| , \end{aligned}$$

so daß

$$\frac{\partial}{\partial \mathbf{n_y}} \, ln|\mathbf{x}-\mathbf{y}|ds_{\mathbf{y}} = d\omega_{\mathbf{y}} \ .$$

Es folgt, daß die Gleichung (5.11) auch folgendermaßen geschrieben werden kann:

$$f(x) = \pi\mu(x) + \int_{\Gamma} \mu(y)\frac{\partial}{\partial \mathbf{n_y}} \, ln|\mathbf{x}-\mathbf{y}|ds_{\mathbf{y}} \quad \text{für} \quad x \in \Gamma \ . \tag{7.14}$$

7.2.2 Die Integralgleichung für eine einfache Schicht

Gesucht ist eine Lösung u des Dirichlet-Problems

$$u_{xx} + u_{yy} = 0 \ , \quad \text{in } \Omega \tag{7.15}$$
$$u = f \ , \quad \text{auf } \Gamma \ . \tag{7.16}$$

Für festes $\mathbf{y} \in \mathbb{R}^2$ ist

$$u(\mathbf{x}) := ln|\mathbf{x}-\mathbf{y}| = ln\left[(x_1-y_1)^2 + (x_2-y_2)^2\right]^{1/2}$$

eine Lösung (5.15). Deshalb ist

$$u(\mathbf{x}) := \int_{\Gamma} \sigma(\mathbf{y}) ln|\mathbf{y}-\mathbf{x}|ds_{\mathbf{y}} \ , \ x \in \Omega \tag{7.17}$$

eine Lösung Gleichung (5.15) für jedes $\sigma \in C(\Gamma)$. Hierbei ist $s_{\mathbf{y}}$ die Länge entlang Γ.

Damit u die Randbedingung (5.16) erfüllt, ist es erforderlich, daß σ folgende Integralgleichung erfüllt:

$$f(\mathbf{x}) = \int_{\Gamma} \sigma(\mathbf{y}) ln|\mathbf{x}-\mathbf{y}|ds_{\mathbf{y}} \ , \ x \in \Gamma \tag{7.18}$$

Gleichung (5.18) heißt eine *Integralgleichung Erster Art.*

7.2.3 Anwendung der Greenschen Formel

Für $u, v \in C^2(\bar{G})$ gilt:

$$\int\int_{G} (u\Delta v - v\Delta u)dx = \int_{\partial G} \left(u\,\frac{\partial v}{\partial n} - v\,\frac{\partial u}{\partial n}\right) ds \ . \tag{7.19}$$

Für vorgeschriebene $\mathbf{x} \in \Omega$ und $\epsilon > 0$ seien

$$\begin{aligned} v(\mathbf{y}) &:= ln|\mathbf{x}-\mathbf{y}| \quad \text{und} \;\; \Gamma_\epsilon := \{\mathbf{y} : |\mathbf{x}-\mathbf{y}| = \epsilon\} \;, \\ G &:= \Omega \backslash B_\epsilon \;,\; B_\epsilon := \{\mathbf{y} : |\mathbf{x}-\mathbf{y}| \le \epsilon\} \;. \end{aligned}$$

Dann gilt, falls $\Delta u = \Delta v = 0$ in Ω :

$$\begin{aligned} &\int\limits_\Gamma \left[u(\mathbf{y}) \frac{\partial v}{\partial \mathbf{n_y}} - v(\mathbf{y}) \frac{\partial u}{\partial \mathbf{n_y}} \right] ds_\mathbf{y} + \\ + &\int\limits_{\Gamma_\epsilon} \left[u(\mathbf{y}) \frac{\partial v}{\partial \mathbf{n_y}} - v(\mathbf{y}) \frac{\partial u}{\partial \mathbf{n_y}} \right] ds_\mathbf{y} = 0 \;. \end{aligned} \tag{7.20}$$

Durch einen Grenzwertprozeßfolgt:

$$-2\pi u(\mathbf{x}) + \int\limits_\Gamma \left[u \frac{\partial}{\partial \mathbf{n_y}} \, ln|\mathbf{x}-\mathbf{y}| - \frac{\partial u}{\partial \mathbf{n_y}} \, ln|\mathbf{x}-\mathbf{y}| \right] ds_\mathbf{y} \;,\; \mathbf{x} \notin \Gamma \tag{7.21}$$

und deshalb

$$\pi u(\mathbf{x}) = \int\limits_\Gamma \left[f(\mathbf{y}) \frac{\partial}{\partial \mathbf{n_y}} \, ln|\mathbf{x} - -\mathbf{y}| - \frac{\partial u}{\partial \mathbf{n_y}} \, ln|\mathbf{x}-\mathbf{y}| \right] ds_\mathbf{y} \;,\; \mathbf{x} \in \Gamma \;. \tag{7.22}$$

Es ist hiermit möglich, die gemischte Randwertaufgabe

$$\alpha u + \beta \frac{\partial u}{\partial n} = \gamma \;,\; \mathbf{x} \in \Gamma$$

zu lösen.

7.3 Numerische Lösung von Integralgleichungen

Im vorherigen Absatz haben wir u.a. folgende Integralgleichung betrachtet:

$$\mu(\mathbf{x}) + \frac{1}{\pi} \int\limits_{\Gamma} \mu(\mathbf{y}) \, \frac{\partial}{\partial \mathbf{n_y}} \, ln|\mathbf{x} - \mathbf{y}| ds_{\mathbf{y}} = f(\mathbf{x}) \; , \; \mathbf{x} \in \Gamma \; . \tag{7.23}$$

Dies ist ein spezieller Fall des folgenden Problems: X sei ein Banachraum.

$K \in B(X) :=$ { Menge stetiger beschränkter linearer Abbildungen von X in sich }.

$f \in X$ sei gegeben. Gesucht ist $x \in X$:

$$(I - K)x = f \; . \tag{7.24}$$

Naheliegend ist der Ansatz, den Operator K durch eine Folge von Operatoren $K_n \in B(X)$, $n \in \mathbb{N}$ und f durch $\{f_n \in X_n \subset X\}_{n \in \mathbb{N}}$ zu approximieren und x durch die Lösung der Gleichung

$$(I - K_n)x_n = f_n \tag{7.25}$$

zu approximieren.

Satz 7.1 *X sei ein Banachraum, $A \in B(X)$, mit $\|A\| < 1$. Dann gilt:*

$$\begin{aligned} (I-A)^{-1} &\in B(X) \\ \|(I-A)^{-1}\| &\leq \frac{1}{1 - \|A\|} \; . \end{aligned}$$

<u>Beweis:</u>
Die Abbildung $(I - A)$ ist injektiv, weil $(I - A)x_1 = (I - A)x_2$ impliziert

$$\|x_1 - x_2\| = \|A(x_1 - x_2)\| < \|x_1 - x_2\| \; .$$

Für $n \in \mathbb{N}$ sei

$$\begin{aligned} T_n &:= I + A + \ldots + A^n \\ T_n &: X \to X \; . \end{aligned}$$

Es folgt, daß

$$\|T_n\| \leq \frac{1}{1 - \|A\|} .$$

Für $x \in X$ ist $\{T_n x\}_{n\in\mathbb{N}}$ eine Cauchy-Folge in X, weil

$$\|T_n x - T_m x\| \leq \|A\|^{n+1}/(1 - \|A\|) , \quad \text{für } \; m > n .$$

Da X ein Banachraum ist, konvergiert $T_n x$. Wir setzen

$$Tx := \lim_{n\to\infty} T_n x .$$

Es läßt sich leicht zeigen, daß $T \in B(X)$. Da

$$(I - A)T_n x = T_n(I - A)x = (I - A^{n+1})x ,$$

folgt

$$(I - A)Tx = T(I - A)x = x ,$$

so daß

$$T = (I - A)^{-1} .$$

Satz 7.2 *X sei ein Banachraum, $A, B \in B(X)$. Existiere $A^{-1} \in B(X)$. Wenn*

$$\|A - B\| < \frac{1}{\|A^{-1}\|} ,$$

dann gilt: $B^{-1} \in B(X)$ und

$$\begin{aligned} \|B^{-1}\| &\leq \frac{\|A^{-1}\|}{1 - \|A^{-1}\| \cdot \|A - B\|} \\ \|B^{-1} - A^{-1}\| &\leq \|A^{-1}\| \cdot \|A - B\| \cdot \|B^{-1}\| . \end{aligned}$$

Beweis:

$$B = A - (A - B) = A(I - C) ,$$

wobei

$$C \in B(X) , \; C := A^{-1}(A - B) .$$

Da $\|C\| < 1$, folgt aus Satz 5.1, daß $(I - C)^{-1} \in B(X)$ und

$$\|(I - C)^{-1}\| \leq \frac{1}{1 - \|C\|} \leq \frac{1}{1 - \|A^{-1}\| \cdot \|A - B\|} .$$

Es folgt, daß

$$B^{-1} = (I - C)^{-1} A^{-1} \in B(X) \, .$$

Die Abschätzung für $\|B^{-1} - A^{-1}\|$ folgt aus der Gleichung

$$B^{-1} - A^{-1} = B^{-1}(A - B)A^{-1} \, .$$

Satz 7.3 *$\{K_n\}_{n\in\mathbb{N}}$ sei eine Folge in $B(X)$ mit $\|K - K_n\| \to 0$ für $n \to \infty$. Ferner existiere $(I - K)^{-1} \in B(X)$. Dann gilt:*

1. *Es existiere $n_0 \in \mathbb{N}$, so daß für $n \geq n_0$*

$$(I - K_n)^{-1} \in B(X) \, .$$

2.

$$\sup\{\|(I - K_n)^{-1}\| : n \geq n_0\} < \infty$$

3.

$$\|(I - K)^{-1} - (I - K_n)^{-1}\| \to 0 \quad \textit{mit } n \to \infty$$

4.

$$\begin{aligned} \|(I - K_n)^{-1}\| &\leq \frac{\|(I - K)^{-1}\|}{1 - \|(I - K)^{-1}\| \cdot \|K - K_n\|} \\ \|(I - K)^{-1}\| &\leq \frac{\|(I - K_n)^{-1}\|}{1 - \|(I - K_n)^{-1}\| \cdot \|K - K_n\|} \end{aligned}$$

Seien $\{f_n\}_{n\in\mathbb{N}}$ Folge in X mit $f_n \to f \in X$, x die Lösung von (5.24) und für $n \geq n_0$ x_n die Lösung von (5.25).

Dann gilt:

5.

$$\begin{aligned} \|x - x_n\| &\leq \|(I - K_n)^{-1}\| \left(\|f - f_n\| + \|K - K_n\| \, \|x\|\right) \to 0 \\ \|x - x_n\| &\leq \|(I - K)^{-1}\| \left(\|f - f_n\| + \|K - K_n\| \, \|x_n\|\right) \to 0 \end{aligned}$$

Beweis:
1 - 4 folgen aus Satz 5.2.5 folgt aus der Identität:

$$\begin{aligned} f - f_n &= (I - K)x - (I - K_n)x_n \\ &= (I - K_n)(x - x_n) + (K_n - K)x \, . \end{aligned}$$

Satz 7.4 *: Banach-Steinhaus (Spezialfall)*
Seien X ein Banachraum, $\{S_n\}$ eine Folge Operatoren $S_n \in B(X)$. Dann gilt 1 $\Leftrightarrow$ 2 + 3, wobei

1. $\lim\limits_{n\to\infty} S_n x$ *existiere für jedes* $x \in X$
2. $\lim\limits_{n\to\infty} S_n x$ *existiere für jedes* $x \in Y$ *mit* $\bar{Y} = X$
3. $\sup\limits_n \|S_n\| < \infty$.

Definition 7.1 *S sei eine Untermenge eines Banachraums X. S ist präkompakt, falls für jedes $\epsilon > 0$ eine endliche Anzahl von Elementen $y_k \in X$ existiert, so daß für jedes $y \in S$*

$$\inf_k \|y - y_k\| \leq \epsilon \; .$$

D.h. die Menge der Kugeln S_k mit Radius ϵ und Mittelpunkte y_k enthält S. Die Menge $\{y_k\}$ heißt ϵ-Netz für S.

Definition 7.2 *$K : X \to X$ heißt kompakt, falls KS präkompakt ist für jede beschränkte Menge $S \subset X$. D.h. K bildet beschränkte Mengen in präkompakte Mengen ab.*

Satz 7.5 *Seien X ein Banachraum, $\{S_n\}_{n\in\mathbb{N}}$ eine Folge Operatoren, $S_n \in B(X)$, $A \subset X$ präkompakt,*

$$S_n x \to Sx \in X \quad \text{mit } n \to \infty \quad \text{für jedes } x \in X \; .$$

Dann gilt: für jedes $\epsilon > 0$ existiere N :

$$\|S_n x - Sx\| \leq \epsilon \; , \quad \text{für } n \geq N \quad \text{und alle } x \in A \; .$$

D.h. die Konvergenz der Folge $\{S_n x\}$ ist gleichmäßig für $x \in A$.

Beweis:
Wähle $\epsilon > 0$. Da die Menge A präkompakt ist, gibt es ein ϵ-Netz für A :

$$x_1, \ldots, x_m \; .$$

Sei $Sx = \lim S_n x$. Es folgt (Banach-Steinhaus), daß es eine Konstante M gibt, mit

$$\|S\| \leq M \; , \quad \|S_n\| \leq M \quad \text{für } n \in \mathbb{N} \; .$$

Wähle N mit

$$\|Sx_k - S_n x_k\| \leq \epsilon \quad \text{für } 1 \leq k \leq m \ , \quad n \geq N \ .$$

Sei nun $x \in A$. Da $\{x_k\}$ ein ϵ-Netz ist, gibt es ein x_k mit

$$\|x - x_k\| \leq \epsilon \ .$$

Für $n \geq N$ folgt, daß

$$\begin{aligned} \|Sx - S_n x\| &\leq \|S_n x_k - Sx_k\| + \|(S - S_n)(x - x_k)\| \ , \\ &\leq \epsilon + 2M\epsilon \ . \end{aligned}$$

Das Projektionsverfahren

$P \in B(X)$ sei eine Projektion (d.h. $P^2 = P$) mit $\tilde{X} := PX$ und dim $\tilde{X} < \infty$.
Als Approximation zur Lösung x der Gleichung (5.24) nehme man $\tilde{x} \in X$ mit

$$P(I - K)\tilde{x} = Pf \ , \tag{7.26}$$

oder

$$(I - PK)\tilde{x} = Pf \ . \tag{7.27}$$

Es folgt:

$$\tilde{x} \in X \ . \tag{7.28}$$

Benutzt man eine Folge von endlichdimensionalen Unterräumen $\{X_n\}_{n\in\mathbb{N}}$ und Projektionen $\{P_n\}_{n\in\mathbb{N}}$ mit jeweils $P_n X = X_n$, erhalten wir Approximationen x_n mit

$$(I - P_n K)x_n = P_n f \ , \quad x_n \in X_n \ . \tag{7.29}$$

Satz 7.6 *Seien X ein Banachraum, $K \in B(X)$ kompakt, $\{P_n\}$ eine Folge Projektionen, $(I-K)^{-1} \in B(X)$, $P_n x \to x$ für jedes $x \in X$, $K_n := P_n K$.
Dann gilt: $\|K_n - K\| \to 0$ mit $n \to \infty$.*

Es gibt $N \in \mathbb{N}$ derart, daß $(I - K_n)^{-1} \in B(X)$ für $n \geq N$. Sei $n \geq N$,

$$(I - K_n)x_n = P_n f \ , \ (I - K)x = f \ .$$

Dann gilt:

$$\|x - x_n\| \longrightarrow 0 \quad \text{für } n \to \infty \, .$$

<u>Beweis:</u>
Sei B die Einheitskugel um X. Dann ist $A := K(B)$ präkompakt. Es folgt:

$$\begin{aligned} \|K - P_n K\| &= \sup_{x \in B} \|Kx - P_n Kx\| \, , \\ &= \sup_{y \in A} \|y - P_n y\| \longrightarrow 0 \, , \quad \text{für } n \to \infty \, , \end{aligned}$$

wie aus Satz 5.7 folgt.

Wegen des Banach-Steinhaus Satzes existieren Konstanten $M \in \mathbb{R}$ und $N \in \mathbb{N}$ derart, daß für $n \geq N$

$$\begin{aligned} (I - K_n)^{-1} &\in B(X) \, , \\ \|(I - K_n)^{-1}\| &\leq M \, . \end{aligned}$$

Sei, für $n \geq N$,

$$(I - K)u = f \, , \quad (I - K_n)u_n = f_n = P_n f \, .$$

Dann folgt:

$$\|u - u_n\| \leq \|(I - K)^{-1}(K - K_n)u_n\| + \|(I - K)^{-1}\| \cdot \|P_n f - f\| \, ,$$

so daß

$$\|u - u_n\| \longrightarrow 0 \quad \text{für } n \longrightarrow \infty \, .$$

Satz 7.7 *Es gelte die Voraussetzung von Satz 5.8.*
Falls $\|K_n\| = \|P_n K\| < 1$, erhalten wir:

$$\begin{aligned} Q_n &:= I - P_n \\ (I - P_n K)(u - u_n) &= Q_n(u - x) \, , \quad \forall x \in X_n \, , \end{aligned}$$

woraus folgt:

$$\|u - u_n\| \leq \frac{1}{1 - \|K_n\|} \cdot (1 + \|P_n\|) \inf_{x \in X_n} \|u - x\| \, .$$

Diese Gleichung besagt, daß die Lösung u_n optimal ist.

7.4 Anwendungen

Wir betrachten die Gleichung

$$x(s) - \int_a^b k(s,t)x(t)dt = f(s)\ , \quad a \le s \le b\ , \tag{7.30}$$

d.h.

$$(I-K)x = f \tag{7.31}$$

im Banachraum

$$X := C[a,b]\ .$$

Satz 7.8 *Sei* $k : [a,b] \times [a,b] \to \mathbb{R}$ *stetig. Sei* $K : C[a,b] \to C[a,b]$*,*

$$(Kx)(s) = \int_a^b k(s,t)x(t)dt\ .$$

Dann ist K *ein kompakter Operator.*

<u>Beweis:</u>
(siehe auch Hackbusch:*Integralgleichungen*, Teubner Verlag, Satz 3.2.6)
k ist stetig auf einer kompakten Teilmenge von $\mathbb{R}^2$ und deshalb auch gleichmäßig stetig. Wähle $\epsilon > 0$. Es gibt $\delta > 0$ mit

$$|k(s_1,t) - k(s_2,t)| \le \epsilon \quad \text{für alle} \ \ t, s_1, s_2 \quad \text{mit} \ \ |s_1 - s_2| \le \delta\ .$$

Sei

$$a = s_0 < s_1 < \cdots < s_n = b$$

mit $|s_j - s_{j-1}| \le \delta$. Sei B die Einheitskugel um X,

$$B := \{x \in X : \|x\| \le 1\}\ ,$$

und

$$M := |b-a| \cdot \max_{\substack{a \le s \le b \\ a \le t \le b}} |k(s,t)|\ .$$

Sei $S := K(B)$. Es gilt:

$$\|y\| \le M \quad \text{für alle} \ \ y \in S\ .$$

Für jedes $y \in S$ gilt auch:

$$\begin{aligned} |y(s) - y(s_\ell)| &\leq \int\limits_a^b |k(s,t) - k(s_\ell,t)| \cdot |y(t)| dt \\ &\leq \epsilon \cdot (b-a) \quad \text{für } s \in [s_\ell, s_{\ell+1}] \ . \end{aligned}$$

Teile das Intervall $[-M, +M]$ in m Teilintervalle mit jeweils der Länge kleiner als ϵ,

$$[-M, +M] = \bigcup_{j=1}^{m} I_j \ .$$

Sei

$$S_{k_0 \dots k_n} := \{y \in S : y(s_i) \in S_{k_i} \ , \ 0 \leq i \leq n\} \ .$$

Falls $S_{k_0 \dots k_n} \neq \emptyset$ wähle *ein* Element $\tilde{y} \in S_{k_0 \dots k_n}$. Es gibt insgesamt nicht mehr als m^{n+1} Elemente $\tilde{y}$. Sei $\tilde{S}$ die Menge aller $\tilde{y}$.
Sei nun $y \in S$. Es existiert $\tilde{y} \in \tilde{S}$ mit:

$$|\tilde{y}(s_\ell) - y(s_\ell)| \leq \epsilon \ , \ 0 \leq \ell \leq n \ .$$

Es folgt:

$$\|\tilde{y} - y\| \leq \epsilon + 2\epsilon \cdot (b-a) \ .$$

Damit ist bewiesen, daß $\tilde{S}$ ein ϵ-Netz für S ist, d.h. S ist präkompakt.

Es gibt natürlich mehrere Möglichkeiten, die Projektionen P_n zu wählen. Hier wird nur eine Möglichkeit kurz erläutert. Sei

$$\begin{aligned} s_i &= a + \frac{b-a}{n} \cdot i \ , \ 0 \leq i \leq n \ , \\ (P_n x)(s) &= \frac{(s - s_i) \cdot x(s_{i+1}) + (s_{i+1} - s) \cdot x(s_i)}{s_{i+1} - s_i} \ , \ s_i \leq s \leq s_{i+1} \end{aligned}$$

Um die Lösung x_n der Gleichung

$$(I - P_n K) x_n = P_n f$$

zu bestimmen, genügt es, die Werte von x_n in den Stützpunkten s_i zu berechnen. Es folgt:

$$(I_n - A_n) \mathbf{x} = \mathbf{b}_n \ ,$$

mit

$$
\begin{aligned}
\mathbf{x}_n &= (x_n(s_0), \ldots, x_n(s_n)) \;, \\
\mathbf{b}_n &= (f(s_0), \ldots, f(s_n)) \;, \\
A_n &= \left(a_{ij}^{(n)}\right) \;, \\
a_{i,j}^{(n)} &= \int\limits_{s_{j-1}}^{s_j} k(s_i, t) \cdot \frac{t - s_{j-1}}{s_j - s_{j-1}} \cdot dt + \\
&+ \int\limits_{s_j}^{s_{j+1}} k(s_i, t) \cdot \frac{s_{j+1} - t}{s_{j+1} - s_j} \cdot dt \;,
\end{aligned}
$$

wobei das erste bzw. letzte Integral für $j = 0$ bzw. $j = n$ wegfällt.

Kapitel 8

Finite Elemente: Theoretische Vorbereitungen

8.1 Einführung

Die Methode der finiten Elemente unterscheidet sich vom Differenzenverfahren dadurch:

1. Die Approximationen u_h zur gesuchten Lösung u sind "echte" Funktionen und nicht nur Gitterfunktionen.

2. Die Bedingungen, die durch die Approximationen erfüllt werden müssen, werden durch Variationsprinzipien oder Integralbedingungen hergeleitet.

3. Die Approximationen sind Summen von Funktionen, die nur begrenzt oft differenzierbar sind, dafür aber kleine Träger besitzen.

Um die erforderlichen Vorbereitungen zu motivieren, betrachten wir zuerst ein sehr einfaches Beispiel.

Beispiel 8.1 Gesucht ist die Lösung der Randwertaufgabe

$$-u''(x)+u(x)=f(x)\,,\quad 0<x<1\,, \tag{8.1}$$

$$u(0)=0, u(1)=0 \tag{8.2}$$

mit

$$u\in C^2(0,1)\cap C[0,1]\,.$$

Sei $\varphi\in C_0^\infty(0,1)$, d.h. φ ist unendlich oft differenzierbar und hat einen kompakten Träger in $(0,1)$. Wird die Gleichung (8.1) mit φ multipliziert und anschließend integriert, so folgt nach partieller Integration, daß

$$\int_0^1 [u'(x)\varphi'(x) + u(x)\varphi(x)]dx = \int_0^1 \varphi(x)f(x)dx \ , \tag{8.3}$$

oder

$$a(u,\varphi) = < b, \varphi > \quad \text{für alle} \ \ \varphi \in C_0^\infty(0,1) \ , \tag{8.4}$$

wobei

$$\begin{aligned} a(u,v) &:= \int_0^1 [u'(x)v'(x) + u(x)v(x)]dx \ , \\ < b, \varphi > &:= \int_0^1 f(x)\varphi(x)dx \ .. \end{aligned}$$

Die Abbildungen a und b sind nur für stetig differenzierbare Funktionen bzw. stetige Funktionen definiert. Wir setzen jetzt voraus, daß die Definitionsbereiche erweitert werden können.

Sei nun X ein Hilbert-Raum mit

$$X \supset C^2(0,1) \cap C[0,1] \ .$$

Sei a und b eine Bilinearform

$$a : X \times X \longrightarrow \mathbb{R} \tag{8.5}$$

bzw. ein lineares Funktional

$$b : X \longrightarrow \mathbb{R} \ . \tag{8.6}$$

Gesucht wird ein Element $u \in X$ mit

$$a(u,v) = < b, v > \ , \quad \text{für alle} \ \ v \in X \ . \tag{8.7}$$

Zur Lösung von (8.7) kann man das Galerkin-Verfahren anwenden:

1. Wähle

$$X_n = \text{span}\{x_1, \ldots, x_n\} \subset X \ . \tag{8.8}$$

2. Bestimme $u_n \in X_n$ mit

$$a(u_n, v_n) =< b, v_n > \quad \text{für alle } \; v_n \in X_n \; . \tag{8.9}$$

Diese Lösungsmethode ist klassisch. Die Finite-Elemente-Methode besteht darin, die Basisfunktionen x_k so zu wählen, daß

1. jedes x_k einen kleinen kompakten Träger besitzt,
2. x_k stückweise glatt ist,
3. x_k durch einfache Funktionen, z.b. Polynome, definiert ist.

Fragen, die jetzt zu beantworten sind, sind folgende:

1. Existiert die Lösung u von (8.7)?
2. Existiert die Lösung u_n von (8.9)?
3. Wie kann ein geeigneter Hilbert-Raum X definiert werden?
4. Wie können geeignete Basisfunktionen definiert werden?
5. Gilt $||u - u_n|| \longrightarrow 0$ für $n \longrightarrow \infty$?
6. Kann $||u - u_n||$ abgeschätzt werden?
7. Welche numerischen Aufgaben müssen gelöst werden?

Die ersten beiden Fragen (Existenz und Eindeutigkeit der Lösungen u und u_n) werden in diesem Kapitel beantwortet. Die Konstruktion des Raumes X führt unausweichlich zu einer Besprechung über Sobolewsche Räume, die im nächsten Kapitel durchgeführt wird. Erst danach können die üblichen Fragen behandelt werden.

8.2 Variationsgleichungen im Hilbertraum

Ein reeller Hilbertraum H ist ein Vektorraum über $\mathbb{R}$ mit Skalarprodukt $(\cdot, \cdot)$, der bezüglich der Norm, die durch das Skalarprodukt induziert wird, vollständig ist. Es gilt:

1. $(u, u) \geq 0$, für alle $u \in H$, $(u, u) = 0$ nur für $u = 0$.
2. $(u, v) = (v, u)$, für alle $u, v \in H$.
3. $(\alpha u_1 + \beta u_2, v) = \alpha(u_1, v) + \beta(u_2, v)$ für alle $\alpha, \beta \in \mathbb{R}$ und $u_1, u_2, v \in H$.

Ein Hilbertraum ist also auch ein Banachraum mit der Norm

$$||u|| := (u,u)^{1/2}$$

und es gilt:

1. $||u+v|| \leq ||u|| + ||v||$, (Dreiecksungleichung)
2. $|(u,v)| \leq ||u|| \cdot ||v||$, (Schwarzsche Ungleichungen)

   Beweis: Man betrachtet $F(\lambda) := (u-\lambda v, u-\lambda v) \geq 0$ an der Stelle, wo $F'(\lambda) = 0$.

3. $||u+v||^2 + ||u-v||^2 = 2(||u||^2 + ||v||^2)$, (Parallelogrammidentität)

Eine lineare Abbildung F von einem Banachraum X nach $\mathbb{R}^1$ heißt ein lineares Funktional. Ein Funktional F heißt beschränkt, wenn es eine Konstante C gibt, so daß

$$||Fx|| \leq C||x||$$

und

$$||F|| := \sup_{\substack{x \in X \\ x \neq 0}} \frac{||Fx||}{||x||}$$

heißt die Norm von F.
Eine Abbildung

$$a : X \times X \longrightarrow \mathbb{R}$$

heißt eine Bilinearform (bzw. Sesquilinearform in $\mathbb{C}$), wenn gilt:

$$\begin{aligned} a(\alpha_1 u_1 + \alpha_2 u_2, v) &= \alpha_1 a(u_1, v) + \alpha_2 a(u_2, v) \\ a(u, \beta_1 v_1 + \beta_2 v_2) &= \beta_1 a(u, v_1) + \beta_2 a(u, v_2) \ . \end{aligned}$$

a heißt beschränkt, wenn

$$a(u,v) \leq C||u|| \cdot ||v|| \ , \quad \text{für alle } \ u, v \in X$$

und

$$||a|| := \sup \frac{a(u,v)}{||u|| \cdot ||v||} \ .$$

a heißt koerziv (X-elliptisch), falls es eine Konstante $\alpha > 0$ gibt, so daß

$$a(u,u) \geq \alpha \|u\|^2 \ , \quad \text{für alle} \ \ u \in X \ .$$

Sei M eine Teilmenge des Hilbertraums H.

$$M^{\perp} := \{u \in H : (u,m) = 0 \quad \text{für alle} \ \ m \in M\} \ .$$

Die folgenden Sätze sind in der Funktionalanalysis gut bekannt. Siehe z.B. Gilbarg, Trudinger [1977], S. 76.

Satz 8.1 (Projektionssatz) *Sei M ein abgeschlossener Teilraum des Hilbertraums H. Für jedes $u \in H$ gilt $u = v + w$, wobei $v \in M$ und $w \in M^{\perp}$.*

Beweis: Ist $u \in M$, setze man $v = u$ und $w = 0$.
Ist $u \notin M$, sei

$$d := \inf_{v \in M} \|u - v\| \ .$$

Sei $\{v_n\}$ eine minimierende Folge, d.h. $\|u - v_n\| \longrightarrow d$ für $n \longrightarrow \infty$. Es gilt das Parallelogrammgesetz:

$$\|x + y\|^2 + \|x - y\|^2 = 2(\|x\|^2 + \|y\|^2) \ .$$

Mit $x = u - v_n$, $y = u - v_m$ folgt

$$4 \left\| u - \frac{v_n + v_m}{2} \right\|^2 + \|v_n - v_m\|^2 = 2\|u - v_n\|^2 + 2\|u - v_m\|^2 \ . \quad (*)$$

Sei $\epsilon > 0$. Es gibt N mit

$$\|u - v_k\|^2 \leq d^2 + \epsilon \ , \quad \text{für alle} \ \ k \geq N \ .$$

Wähle $n, m \geq N$. Es folgt aus $(*)$, daß

$$\begin{aligned} \|v_n - v_m\|^2 &\leq 2(\|u - v_n\|^2 + \|u - v_m\|^2, -4 \left\| u - \frac{v_n + v_m}{2} \right\|^2 \\ &\leq 2(d^2 + \epsilon + d^2 + \epsilon) - 4d^2 \\ &= 4\epsilon \ . \end{aligned}$$

D.h. $\{v_n\}$ ist eine Cauchy-Folge. Da H ein Hilbert-Raum und M abgeschlossen ist, gibt es $v \in M$ mit

$$v_n \longrightarrow v \quad \text{für} \ \ n \longrightarrow \infty$$

und

$$\|v - u\| = d \, .$$

Sei nun $w := u - v$,

$$u = v + w \, .$$

Sei $z \in M$ und $\alpha > 0$. Dann ist $v + \alpha z \in M$. Folglich

$$\begin{aligned} d^2 &\leq \|u - (v + \alpha z)\|^2 \\ &= \|w\|^2 - 2\alpha(w, z) + \alpha^2 \|z\|^2 \\ &= d^2 - 2\alpha(w, z) + \alpha^2 \|z\|^2 \end{aligned}$$

D.h.

$$(w, z) \leq \frac{\alpha}{2} \|z\|^2 \quad \text{für alle} \quad \alpha, z$$

oder

$$(w, z) = 0 \, , \quad \forall z \, .$$

Satz 8.2 (Riesz) *Jedes lineare beschränkte Funktional F auf dem Hilbertraum H ist von der Form*

$$F(x) = (x, y) \qquad\qquad (*)$$

wobei $y \in H$ durch F eindeutig bestimmt ist. Umgekehrt definiert $(*)$ für jedes $y \in H$ ein beschränktes lineares Funktional auf H. Es gilt $\|F\| = \|y\|$.

Beweis: Sei

$$M = \{x : F(x) = 0\} \, .$$

M ist ein abgeschlossener Teilraum von H. Ist $M = H$, so erfüllt $y = 0$ die Voraussetzungen des Satzes.

Sei $M \neq H$ und $x \in H \backslash M$. Nach dem Projektionssatz gilt

$$x = y + z$$

mit $y \in M$ und $z \in M^{\perp}$.

Da $x \notin M$, ist $z \neq 0$.

Sei nun $x \in H$ beliebig. Man setze

$$w := x - \frac{F(x)}{F(z)} \cdot z \ .$$

Es gilt $F(w) = 0$, so daß $w \in M$ und $w \perp z$. D.h.

$$(w, z) = 0 = (x, z) - \frac{F(x)}{F(z)} \cdot \|z\|^2 = 0 \ .$$

Durch Umformung folgt:

$$F(x) = (x, y)$$

mit

$$y = \frac{F(z)}{\|z\|^2}\, z \ .$$

Es ist leicht einzusehen, daß y eindeutig ist.
Zuletzt folgt aus

$$\|F\| = \sup_x \frac{|(x, y)|}{\|x\|} \leq \frac{\|x\| \cdot \|y\|}{\|x\|} = \|y\|$$

und

$$\|F\| = \sup_x \frac{|(x, y)|}{\|x\|} \geq \frac{|(y, y)|}{\|y\|} = \|y\| \ ,$$

daß $\|F\| = \|y\|$.

Satz 8.3 (Lax-Milgram) *Sei a eine beschränkte, koerzive Bilinearform auf dem Hilbertraum H und F ein beschränktes lineares Funktional auf H. Dann gibt es ein eindeutiges Element $f \in H$, so daß*

$$a(u, f) = F(u) \quad \textit{für alle} \ \ u \in H \ .$$

Beweis: Sei $v \in H$. Die Abbildung

$$u \longmapsto a(u, v)$$

ist ein beschränktes lineares Funktional. Wegen des Rieszschen Satzes gibt es ein $f_v \in H$ mit

$$a(u, v) = (u, f_v) \quad \text{für alle} \ \ u \in H \ .$$

Die Abbildung $T : H \longrightarrow H$,

$$T : v \longrightarrow f_v$$

ist beschränkt und linear,

a) Beschränktheit

$$\|f_v\|^2 = (f_v, f_v) = a(f_v, v) \leq \|a\| \cdot \|f_v\| \cdot \|v\| ,$$

so daß

$$\|f_v\| = \|Tv\| \leq \|a\| \cdot \|v\|$$

b) Linearität

Sei

$$\begin{aligned} a(u, v_1) &= (u, Tv_1) , \\ a(u, v_2) &= (u, Tv_2) \end{aligned}$$

Dann gilt:

$$a(u, v_1 + v_2) = (u, Tv_1 + Tv_2) ,$$

d.h.

$$T(v_1 + v_2) = Tv_1 + Tv_2 .$$

Es gilt auch, daß

$$\begin{aligned} \alpha\|v\|^2 &\leq a(v, v) = (v, Tv) \leq \|v\| \cdot \|Tv\| , \\ &\Rightarrow \|Tv\| \geq \alpha\|v\| . \end{aligned}$$

T ist also injektiv und T^{-1} ist beschränkt. Der Wertebereich $R(T)$ ist abgeschlossen.

Es folgt, daß $R(T) = H$, d.h. T ist surjektiv. Denn sonst gibt es wegen des Projektionssatzes mit $M = R(T)$ ein Element

$$\begin{aligned} z &\in H , \ z \neq 0 , \\ z &\in M^{\perp} . \end{aligned}$$

D.h.

$$a(z,v) = (z,Tv) = 0 \quad \text{für alle } v \in H \ .$$

Insbesondere

$$a(z,z) = (z,Tz) = 0 \ ,$$

so daß

$$z = 0 \ .$$

Die inverse Abbildung T^{-1} ist deshalb eine lineare, beschränkte Abbildung von H auf H.

Es gilt:

$$a(u,T^{-1}w) = (u,w) \quad \text{für alle } u,w \in H \ .$$

Wegen des Rieszschen Satzes gibt es ein Element $g \in H$ mit

$$F(u) = (u,g) \quad \text{für alle } u \in H \ .$$

Es folgt:

$$F(u) = (u,g) = a(u,T^{-1}g) \quad \text{für alle } u \in H \ .$$

Man setze $f := T^{-1}g$.

Bemerkung 8.1 Es gibt auch ein eindeutiges Element $u \in H$ mit

$$a(u,v) = F(v) \quad \text{für alle } v \in H \ .$$

Satz 8.4 *Sei a eine beschränkte, symmetrische, koerzive Bilinearform auf dem Hilbertraum H und $F(v) := (b,v)$ mit $b \in H$,*

$$E(v) := a(v,v) - 2(b,v) \ . \qquad (*)$$

Dann ist $u \in H$ eine Lösung der Gleichung

$$a(u,v) = (b,v) \qquad (**)$$

genau dann, wenn

$$E(u) = \inf_{v \in H} E(v) \qquad (***)$$

Es existiert genau ein $u \in V$, das $(***)$ erfüllt.

Beweis:
$\Rightarrow$ Sei u die Lösung von $(**)$ und $v \in H$. Dann folgt

$$\begin{aligned} E(v) - E(u) &= a(v,v) - a(u,u) - 2(b, v-u) \\ &= a(v-u, v-u) + 2[a(u, v-u) - (b, v-u)] \\ &= a(v-u, v-u) \\ &\geq 0 . \end{aligned}$$

$\Leftarrow$ u erfülle $(***)$. Dann gilt:

$$E(u) \leq E(u + \lambda v) \quad \text{für alle } \lambda \in \mathbb{R} \, , \ v \in H \, .$$

Es folgt

$$\lambda^2 a(v,v) + 2\lambda a(u,v) - 2\lambda(b,v) \geq 0 \, .$$

Für $\lambda \downarrow 0$ erhalten wir aus

$$\lambda a(v,v) + 2a(u,v) - 2(b,v) \geq 0$$

daß

$$a(u,v) - (b,v) \geq 0 \, .$$

Für $\lambda \uparrow 0$ erhalten wir

$$a(u,v) - (b,v) \leq 0 \, .$$

8.3 Variationsungleichungen: Einleitung

Im letzten Absatz wurde gezeigt, daß die Variationsgleichung

$$a(u,v) = (b,v)\ , \quad v \in H$$

auch als Optimierungsproblem

$$\inf_{v \in H} a(v,v) - 2(b,v)$$

gestellt werden kann. Es gibt eine Erweiterung dieser beiden Probleme, die Variationsungleichung, die sowohl theoretische als auch praktische Anwendung findet. In diesem Absatz betrachten wir ein einfaches Beispiel (siehe Elliot und Ockendon [1982, S. 11]).

Man betrachte das Randwertproblem (RWP)

$$\begin{aligned} -\tilde{u}''(x) &= f(x) := +1\ , \quad 0 \leq x \leq 1 \\ \tilde{u}(0) &= \tilde{u}(1) = 0 \end{aligned} \tag{8.10}$$

Die Lösung

$$\tilde{u}(x) := \frac{x(1-x)}{2} \tag{8.11}$$

erfüllt die Variationsgleichung (VG):

$$a(\tilde{u},v) = (f,v)\ , \quad v \in X \tag{8.12}$$

mit

$$a(u,v) := \int_0^1 u_x v_x \, dx \tag{8.13}$$

$$(f,v) := \int_0^1 f(x)v(x)dx \tag{8.14}$$

und die geeignete Wahl des Raumes X. (Z.B. $X = H_0^1(0,1)$.) Diese Lösung $\tilde{u}$ wird in Abbildung 8.1a) graphisch dargestellt.

Es wird jetzt vorausgesetzt, daß das Problem (8.10) ein physikalisches Problem darstellt, und zwar die Biegung eines Stabes (man denke an Splines!). Es sei jetzt ein Hindernis $y = \varphi(x) := c(1-x)$ vorhanden, so daß die Lösung $u(x)$ den Wert $\varphi(x)$ nicht überschreiten darf. Man stellt sich jetzt vor, daß eine mögliche Lösung $u(x)$ wie in Abbildung 8.1b) aussehen würde. Zwischen $x = 0$ und $x = x_0$ liegt der Stab unterhalb des Hindernisses und erfüllt die vorgegebene Differentialgleichung. Zwischen $x = x_0$ und $x = 1$ liegt der Stab auf dem Hindernis.

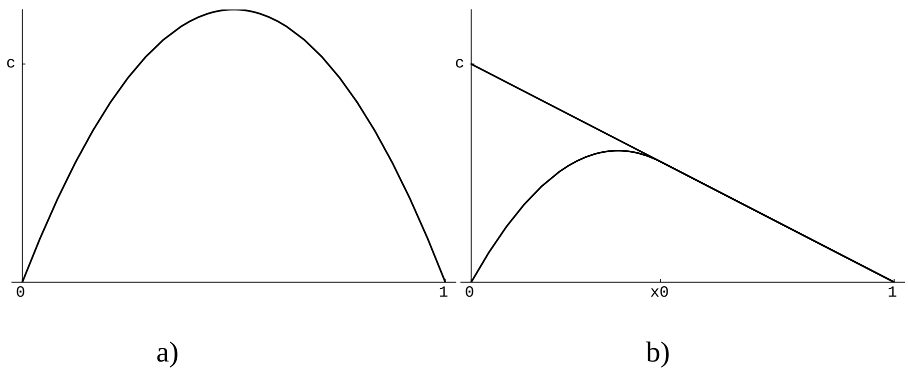

Abbildung 8.1: Elastischer Stab ohne und mit Hindernis

Wenn diese Vorstellung richtig wäre, dann würde u das folgende Problem lösen:

$$\begin{aligned} -u''(x) &= +1\ , \quad 0 \le x \le x_0\ , && (8.15)\\ u(0) &= u(1) = 0\ , \\ u(x) &= \varphi(x)\ , \quad x_0 \le x \le 1\ . && (8.16) \end{aligned}$$

Eine Lösung des Problems (8.15), (8.16) läßt sich leicht berechnen:

$$u(x) = \begin{cases} c(1-x) - \frac{1}{2}(x-x_0)^2 & ,\quad 0 \le x \le x_0 \\ c(1-x) = \varphi(x) & ,\quad x_0 \le x \le 1 \end{cases} \tag{8.17}$$

mit

$$x_0 = \sqrt{2c} \tag{8.18}$$

Die Frage stellt sich nun, ob u eine Lösung der Form (8.12) erfüllt. Sei

$$K = \{v \in X : v(x) \le \varphi(x)\ ,\quad 0 \le x \le 1\}\ .$$

Offensichtlich gilt $u \in K$.

Sei nun $v \in K$.

Es gilt:

$$
\begin{aligned}
a(u, v-u) &= \int_0^1 u_x(v_x - u_x)dx \\
&= \int_0^{x_0} u_x(v_x - u_x)dx + \int_{x_0}^1 u_x(v_x - u_x)dx \\
&= u_x(v-u)]_0^{x_0} + u_x(v-u)]_{x_0}^1 - \int_0^{x_0} u_{xx}(v-u)dx - \int_{x_0}^1 u_{xx}(v-u)dx \\
&= [v(x_0) - \varphi(x_0)][u_x(x_0 - 0) - u_x(x_0 + 0)] + \int_0^{x_0} f \cdot (v-u)dx - \int_{x_0}^1 \varphi_{xx}(v - \varphi)dx \; .
\end{aligned}
$$

Es gilt:

$$
\begin{aligned}
u_x(x_0 - 0) &= u_x(x_0 + 0) = -c \\
\varphi_{xx} &= 0 \; , \quad x_0 \leq x \leq 1 \\
v(x) &\leq \varphi(x) \; , \quad 0 \leq x \leq 1
\end{aligned}
$$

so daß

$$
\begin{aligned}
a(u, v-u) &= + \int_0^{x_0} f \cdot (v-u)dx + 0 \\
&\geq \int_0^{x_0} f \cdot (v-u)dx + \int_{x_0}^1 f \cdot (v-u)dx \\
&= (f, v-u) \; .
\end{aligned}
$$

Zusammenfassend ist u die Lösung der Variationsungleichung (VU).

Finde $u \in K$ mit (VU): $a(u, v-u) \geq (f, v-u)$ für $v \in K$

Es gibt zwei weitere Probleme, die zur VU äquivalent sind:

- Die entsprechende Minimierungsaufgabe heißt - in Anlehnung an Probleme in der Optimierung - das Quadratische Programmierungsproblem (QP):

$$
E(u) = \inf_{v \in K} E(v)
$$

mit

$$E(v) := a(v,v) - 2(f,v) \ .$$

- Das Lineare Komplementaritätsproblem (LKP):

$$\begin{aligned} Au \leq f \quad &, \quad 0 \leq x \leq 1 \\ u \leq \varphi \quad &, \quad 0 \leq x \leq 1 \\ [(Au - f)(x)] \cdot [u(x) - \varphi(x)] = 0 \quad &, \quad 0 \leq x \leq 1 \end{aligned}$$

mit

$$(Au)(x) := -u''(x) \ .$$

Daßu eine Lösung des QP ist, läßt sich leicht bestätigen. Sei

$$a(u, v-u) \geq (f, v-u) \ , \quad \text{für alle } \ v \in K \ .$$

Es folgt, für $v \in K$,

$$\begin{aligned} E(v) - E(u) &= a(v,v) - a(u,u) - 2(f, v-u) \\ &= a(v-u, v-u) + 2[a(u, v-u) - (f, v-u)] \\ &\geq 0 \ . \end{aligned}$$

Daßu eine Lösung des LKP ist, kann ebenfalls leicht bewiesen werden, obwohl die Bedeutung der Gleichungen

$$Au \leq f$$

und

$$u \leq \varphi$$

sorgfältig definiert werden muß.

Der Operator A wird mit Hilfe des Rieszschen Satzes wie folgt definiert:

$$a(u,v) = (Au, v) \ , \quad \text{für alle } \ v \in K \ .$$

Sei

$$K = \varphi - Q \ ,$$

so daß auch folgendermaßen definiert werden kann:

$$Q = \{v \in X : v(x) \geq 0 \quad \text{f.ü.}\}$$

oder

$$Q = \text{Vervollständigung von } \{\Psi \in C_0^\infty(0,1) : \Psi \geq 0\}$$

Die Ungleichung

$$u - \varphi \leq 0$$

wird als $-u + \varphi \in Q$ interpretiert. Die Ungleichung

$$Au \leq f$$

impliziert

$$(Au - f, w) \leq 0 \quad \text{für alle } w \in Q \ .$$

Sei jetzt nun u die Lösung des VU Problems:

$$a(u, v - u) \geq (f, v - u) \quad \text{für alle } v \in K \ ,$$

d.h.

$$(Au - f, v - u) \geq 0 \quad \text{für alle } v \in K \ .$$

Insbesondere ist $v := u - w \in K$ für alle $w \in Q$, so daß

$$(Au - f, w) \leq 0 \quad \text{für alle } w \in Q \ ,$$

d.h., wie behauptet,

$$Au - f \leq 0 \ .$$

Sei $w \in K, w \geq u$. Setzt man $v = w$ in der VU, folgt:

$$(Au - f, w - u) \geq 0 \ .$$

Setzt man $v = 2u - w$ in die VI ein, folgt:

$$(Au - f, u - w) \geq 0 \ .$$

Zusammengenommen folgt:

$$(Au - f, u - w) = 0 \ , \quad \text{für alle } \ w \in K \quad \text{mit } \ w \geq u \ .$$

Betrachte nun eine Folge $\{w_n\} \subset K$, $w_n \geq u$ mit $w_n \longrightarrow \varphi$ in $L^2(0,1)$. Da $Au - f \in \mathcal{L}^2(0,1)$, folgt

$$\begin{aligned} |(Au - f, u - \varphi)| &\leq |(Au - f, u - w_n)| + |(Au - f, w_n - \varphi)| \\ &\leq \|Au - f\|_2 \cdot \|w_n - \varphi\|_2 \ , \end{aligned}$$

also $(Au - f, u - \varphi) = 0$, und daraus folgt die dritte Relation des LKP, da beide Faktoren das Vorzeichen in $(0,1)$ nicht wechseln.

8.3.1 Diskretisierung der Variationsformulierungen

Sei nun

$$0 = x_0 < x_1 < \ldots < x_{n+1} = 1$$

$n+2$ Stützpunkte im Intervall $[0,1]$ mit $x_i = ih, \ h = \frac{1}{n}$ und $S_h \subset X$ die Menge der stetigen stückweisen lineareren Funktionen auf $[0,1]$ mit Stützpunkten $x_0, \ldots x_{n+1}$, so daß die Randbedingungen $s(0) = s(1) = 0$ erfüllt werden. Sei

$$K_h := \{v_h \in S_h : v(x_j) \leq \varphi(x_j) \ , \ \ 0 \leq j \leq n+1\} \ .$$

Wir erhalten die diskrete Variationsungleichung:

Finde $u_h \in K_h$ mit

$$VI_h : a(u_h, v_h - u_h) \geq (f, v_h - u_h) \quad \text{für alle } \ v_h \in K_h \ .$$

Sei u_h die Lösung von VI_h. Man setze

$$\mathbf{u}_h := \begin{pmatrix} u_h(x_1) \\ \vdots \\ u_h(x_n) \end{pmatrix} \in \mathbb{R}^n \ ,$$

$$\varphi_h := \begin{pmatrix} \varphi(x_1) \\ \vdots \\ \varphi(x_n) \end{pmatrix} \in \mathbb{R}^n \ ,$$

$$A_h := \frac{1}{h^2}\begin{pmatrix} 2 & -1 & & O \\ -1 & \ddots & \ddots & \\ & \ddots & \ddots & -1 \\ O & & -1 & 2 \end{pmatrix} \in \mathrm{mat}(n \times n)$$

$$\mathbf{b}_h = \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix} \in \mathbb{R}^n \ .$$

$$\tilde{K}_h := \{\mathbf{v}_h \in \mathbb{R}^n : \mathbf{v}_{h,i} \le \varphi(x_i)\ ,\ \ 1 \le i \le n\}$$

Es folgt

$$\begin{aligned} \widetilde{VU}_h &: \ \mathbf{u}_h^T A_h(\mathbf{v}_h - \mathbf{u}_h) \ge \mathbf{b}_h^T(\mathbf{v}_h - \mathbf{u}_h) \quad \text{für alle } \ \mathbf{v}_h \in \tilde{K}_h \\ \widetilde{QP}_h &: \ E_h(\mathbf{u}_h) \le E_h(\mathbf{v}_h) \quad \text{für alle } \ \mathbf{v}_h \in \tilde{K}_h \end{aligned}$$

mit

$$E_h(\mathbf{v}) = \mathbf{v}^T A_h \mathbf{v} - 2\mathbf{b}_h^T \mathbf{v}_h \quad \text{für alle } \ \mathbf{v}_h \in \mathbb{R}^n$$

und

$$\begin{aligned} \widetilde{LKP}_h : A_h \mathbf{u}_h &\le \mathbf{b}_h \\ \mathbf{u}_h &\le \varphi_h \\ (A_h\mathbf{u}_h - \mathbf{b}_h)^T(\mathbf{u}_h - \varphi_h) &= 0\ , \end{aligned}$$

d.h. falls $\mathbf{u}_{h,i} - \varphi_{h,i} < 0$, dann gilt: $(A_h\mathbf{u}_h - \mathbf{b}_h) = 0$.

Die Lösung $\mathbf{u}_h$ kann durch das projektierte S.O.R.-Verfahren berechnet werden:

$$\begin{aligned} \mathbf{u}_h^{(0)} &\le \varphi_h \quad \text{sonst beliebig} \\ d_{h,i}^{(k)} &= \left(\mathbf{b}_{h,i} - \sum_{j<i} a_{h,ij}\mathbf{u}_{h,j}^{(k+1)} - \sum_{j>i} a_{h,ij}\mathbf{u}_{h,j}^{(k)}\right) / a_{h,ii} - \mathbf{u}_{h,i}^{(k)} \\ \mathbf{u}_{h,i}^{(k+1/2)} &:= \mathbf{u}_{h,i}^{(k)} + \omega d_{h,i}^{(k)} \\ \mathbf{u}_{h,i}^{(k+1)} &= \min\left\{\varphi_{h,i}, \mathbf{u}_{h,i}^{k+1/2}\right\} \end{aligned}$$

mit $\omega \in (0,2)$.

Beispiel 8.2

$$\begin{aligned} c &= \frac{1}{8}\,, \quad h = \frac{1}{3} \\ x_0 &= \sqrt{2c} = \frac{1}{2} \end{aligned}$$

Finde $\mathbf{u}_h = (u_{h,1}, u_{h,2})^T$ mit

$$\begin{aligned} [2u_{h,1} - u_{h,2}] &\leq \frac{1}{9} \\ [2u_{h,2} - u_{h,1}] &\leq \frac{1}{9} \\ u_{h,1} &\leq \varphi(h) = \frac{1-h}{8} = \frac{1}{12} \\ u_{h,2} &\leq \varphi(2h) = \frac{1-2h}{8} = \frac{1}{24}\,. \end{aligned}$$

Die Lösung ist:

$$u_{h,1} = \frac{11}{144}\,, \quad u_{h,2} = \frac{1}{24}\,.$$

Die ersten Gauß-Seidel-Approximationen sind:

k	$u_{h,1}^{(k)}$	$u_{h,2}^{(k)}$
0	0	0
1	0,05555	0,041666
2	0,07388	0,041666

8.4 Variationsungleichungen im Hilbertraum

In diesem Absatz wird die Existenz und Eindeutigkeit der Lösung einer Variationsungleichung bewiesen. Wir folgen dem Beweis in Kinderlehrer und Stampacchia [1980]. Damit erhalten wir auch einen neuen Beweis des Satzes von Lax und Milgram.

Satz 8.5 *Sei $a(u, v)$ eine koerzive Bilinearform auf einem Hilbertraum H mit der Koerzivitätskonstante α, K eine abgeschlossene konvexe Teilmenge von H und f ein lineares beschränktes Funktional auf H (d.h. $f \in H'$). Dann existiert ein eindeutiges Element $u \in K$, so daß*

$$a(u, v-u) \geq (f, v-u) \quad \textit{für alle } \; v \in K \qquad (*)$$

Die Abbildung $f \longrightarrow u$ ist Lipschitz; seien u_1 und u_2 Lösungen des Problems $(*)$ für f_1 bzw. $f_2 \in H'$, dann gilt

$$\|u_1 - u_2\| \leq \frac{1}{\alpha}\|f_1 - f_2\| \qquad (**)$$

Beweis:

Die Ungleichung $(**)$ wird zuerst bewiesen. Seien u_1 und u_2 Lösungen der Gleichung $(*)$:

$$u_i \in K : a(u_i, v - u_i) \geq (f_i, v - u_i) \quad \text{für} \;\; v \in K \;, \;\; i = 1, 2 \;.$$

Setze $v = u_2$ in der Variationsungleichung für u_1 und $v = u_1$ in der Variationsungleichung für u_2 und addiere die beiden Ungleichungen. Es ergibt sich

$$a(u_1 - u_2, u_1 - u_2) \leq (f_1 - f_2 . u_1 - u_2) \;.$$

Unter Benutzung der Koerzivität von a folgt:

$$\begin{aligned} \alpha \|u_1 - u_2\|^2 &\leq \langle f_1 - f_2, u_1 - u_2 \rangle \\ &\leq \|f_1 - f_2\| \cdot \|u_1 - u_2\| \;, \end{aligned}$$

so daß $(**)$ bewiesen ist.

Die Existenz der Lösung u wird in zwei Schritten durchgeführt. Zuerst wird die Existenz bewiesen unter der Annahme, daß die Bilinearform a symmetrisch sei. Anschließend wird der Beweis für den allgemeinen Fall gegeben.

Sei nun vorerst a symmetrisch. Sei

$$I(v) := a(v, v) - 2(f, v) \;, \quad v \in H$$

und

$$d := \inf_{v \in K} I(v) \;.$$

Es gilt

$$\begin{aligned} I(u) &\geq \alpha \|u\|^2 - 2\|f\|_{H'} \cdot \|u\|_H \\ &\geq \alpha \|u\|^2 - (1/\alpha)\|f\|_{H'}^2 - \alpha \|u\|^2 \\ &\geq -(1/\alpha)\|f\|_{H'}^2 \;, \end{aligned}$$

so daß

$$d \geq (1/\alpha)\|f\|_{H'}^2 > -\infty \;.$$

Sei $\{u_n\}$ eine minimierende Folge für I bzgl. K mit

$$\{u_n \in K : d \leq I(u_n) \leq d + (1/n)\} \;.$$

Da a symmetrisch und K konvex ist, folgt

$$\begin{aligned}\alpha\|u_n - u_m\|^2 &\le a(u_n - u_m, u_n - u_m)\\ &= 2a(u_n, u_n) + 2a(u_m, u_m) - 4a\left(\frac{1}{2}(u_n + u_m), \frac{1}{2}(u_n + u_m)\right)\\ &= 2I(u_n) + 2I(u_m) - 4I\left(\frac{1}{2}(u_n + u_m)\right)\\ &\le 2[(1/n) + (1/m)]\ ,\end{aligned}$$

wobei wir zunutze gemacht haben, daß

$$4(f, u_n) + 4(f, u_m) - 8\left(f, \frac{u_n + u_m}{2}\right) = 0\ .$$

Die Folge $\{u_n\}$ ist deshalb eine Cauchy-Folge, und die abgeschlossene Menge K enthält ein Element u, so daß $\|u_n - u\| \longrightarrow 0$. Da die Bilinearform a stetig ist, schließen wir daraus, daß

$$I(u) = d\ .$$

Wir möchten jetzt beweisen, daß u eine Lösung der Variationsungleichung ist.

Sei $v \in K$, $0 \le \epsilon \le 1$. K ist konvex, so daß

$$u + \epsilon(v - u) = (1 - \epsilon)u + \epsilon v \in K$$

und

$$I(u + \epsilon(v - u)) \ge d = I(u)\ ,$$

d.h.

$$2\epsilon a(u, v - u) + \epsilon^2 a(v - u, v - u) - 2\epsilon(f, v - u) \ge 0$$

oder

$$a(u, v - u) \ge (f, v - u) - \frac{1}{2}\epsilon a(v - u, v - u) \quad \text{für alle } \epsilon\ ,\quad 0 \le \epsilon \le 1\ .$$

Man betrachte den Grenzübergang $\epsilon \longrightarrow 0$ und schließe daraus, daß u die Variationsungleichung $(*)$ erfüllt.

Bei den bisherigen Überlegungen wurde vorausgesetzt, daß a symmetrisch ist. Diese Voraussetzung wird jetzt fallengelassen.

Man definiere drei Bilinearformen a_0, b und a_t mit $t \in [0, 1]$:

$$a_t = a_0(u,v) + tb(u,v) \ , \quad 0 \leq t \leq 1 \ ,$$

wo

$$a_0(u,v) = \frac{1}{2}(a(u,v) + a(v,u))$$

und

$$b(u,v) = \frac{1}{2}(a(u,v) - a(v,u)) \ .$$

a_0 und b sind die symmetrischen bzw. unsymmetrischen Teile von a. Man merke:

$$a_1(u,v) = a(u,v)$$

und

$$a_t(v,v) = a(v,v) \ ,$$

so daß a_t koerziv ist mit der Konstante α.

Wir zeigen, daß die Gleichung $(*)$ eine Lösung hat, indem wir eine endliche Folge $0 = t_0 < t_1 < \cdots < t_n = 1$ konstruieren, so daß die Ungleichung

$$u \in K : a_{tk}(u, v-u) \geq (f, v-u) \ , \quad v \in K$$

eine Lösung $u = u_{tk,f}$ für alle $f \in H^{'}$ besitzt. Dies gilt sicherlich für $t_0 = 0$, da a_0 symmetrisch ist. Die Existenz von $t_1, \ldots, t_n$ folgt aus dem nächsten Lemma.

Lemma 8.1 *Die Variationsungleichung*

$$u \in K : a_\tau(u, v-u) \geq (f, v-u) \ , \quad v \in K$$

besitze eine Lösung u für alle $f \in H^{'}$.

Dann ist die Variationsgleichung

$$u \in K : a_t(u, v-u) \geq (f, v-u) \ , \quad v \in K$$

für alle $f \in H^{'}$ und alle t

$$\tau \leq t \leq \tau + t_0$$

lösbar, wo

$$t_0 = \frac{\alpha}{2M}$$

und

$$M = \sup \frac{|b(u,v)|}{||u|| \cdot ||v||} < +\infty \ .$$

Beweis:

Sei $w \in K$. Definiere das lineare Funktional $F_{t,w} : K \longrightarrow \mathbb{R}$ durch

$$(F_{t,w}, v) = (f, v) - (t - \tau)b(w, v)$$

für $\tau \leq t \leq \tau + t_0$. Sei u die Lösung der Variationsungleichung

$$u \in K : a_\tau(u, v - u) \geq (F_{t,w}, v - u) \ , \quad v \in K \qquad (***) \ .$$

Nach den Voraussetzungen existiert u. Daßu eindeutig ist, folgt sofort aus $(**)$. u ist von der Wahl von w und t abhängig. Wir schreiben

$$u = Tw \ .$$

Der Operator T ist definiert und $T : K \longrightarrow K$.

Wir behaupten, daß T kontrahierend ist mit der Konstante $\frac{1}{2}$. Um eben dies zu zeigen, seien $w_1, w_2 \in K$. Sei

$$\begin{aligned} u_1 &= Tw_1 \ , \\ u_2 &= Tw_2 \ . \end{aligned}$$

Wegen $(**)$ erhalten wir mit

$$\begin{aligned} \|Tw_1 - Tw_2\| &= \|u_1 - u_2\| \\ &\leq \frac{1}{\alpha}\|F_{t,w_1} - F_{t,w_2}\| \\ &\leq \frac{1}{\alpha}|t - \tau| \sup_{v \in H} \frac{|b(w_1 - w_2, v)|}{\|v\|} \\ &\leq \frac{1}{\alpha}|t - \tau| M \|w_1 - w_2\| \\ &\leq \frac{t_0}{\alpha} M \|w_1 - w_2\| \\ &= \frac{1}{2}\|w_1 - w_2\| \end{aligned}$$

Da $T : K \longrightarrow K$ kontrahierend ist, folgt aus dem Banachschen Fixpunktsatz, daß es einen Fixpunkt u gibt, $u = Tu$, d.h (siehe $(***)$)

$$\begin{aligned} a_\tau(u, v - u) &\geq (F_{t,u}, v - u) \\ &= (f, v - u) - (t - \tau)b(u, v - u) \end{aligned}$$

oder

$$a_0(u, v-u) + \tau\, b(u, v-u) \geq (f, v-u) - (t-\tau) b(u, v-u)$$

oder

$$a_0(u, v-u) + t\, b(u, v-u) \geq (f, v-u)$$

oder

$$a(u, v-u) \geq (f, v-u) \ .$$

Literatur:

Alt, H.W.: Lineare Funktionalanalysis. 2. verb. Aufl. Springer, 1992.

Elliot, C.M., **Ockendon**, J.R.: Weak and Variational Methods for Moving Boundary Problems. Pitman, Boston, 1982.

Gilbarg, D., **Trudinger**, N.S.: Elliptic Partial Differential Equations of Second Order. Springer 1977.

Heuser, H.: Funktionalanalysis. Teubner, 1986.

Kinderlehrer, D., **Stampacchia**, G.: An Introduction to Variational Inequalities and their Applications. New York, Academic Press, 1980.

Morrey, C.B. Jr.: Multiple Integrals in the Calculus of Variations. Springer, 1966.

Yosida, K.: Functional Analysis, 2nd edition. Springer, 1968.

Kapitel 9

Sobolew-Räume

9.1 Einführung

Im letzen Kapitel wurde die Existenz und Eindeutigkeit der Lösung der Variationsgleichung:

$$a(u,v) = (b,v)\ , \quad v \in H$$

bzw. der Variationsungleichung

$$a(u, v-u) \geq (f, v-u) \quad \text{für alle} \ \ v \in K \qquad\qquad (*)$$

bewiesen. Vorausgesetz war die Existenz des Hilberraumes X. Die Sobolew- Räume sind in vielen Anwendungen geignete Hilberträume.
Spezielle Fälle der Sobolew-Räume wurden auch früher benutzt (siehe z.B. Courant, Friedrichs, Lewy[1928] oder Courant und Friedrichs[1937]), aber wir verdanken Sobolew wertvolle Beiträge insbesondere die Einbettungssätze.
In diesem Kapitel werden die benötigten Begriffe, einschleißlich die Grundzüge des Lebesgueschen Integrals, kurz erläutert.
Es gibt mehrere Bücher, die die Theorie der Sobolewschen Räume ausführlich beschreiben. Besonders empfehlungswert ist das Buch von Adams[1975].

9.2 Die Lebesguesche Theorie

Es gibt mehrere Möglichkeiten, die Lebesguesche Theorie zu entwickeln. Wir benutzen hier die Darstellung von Forster und untersuchen folgende Themen:

1. Integrierbare Funktionen

2. Meßbare Mengen

3. Konvergenzsätze

1. Integrierbare Funktionen

Definition 9.1 *(Oberintegral, Unterintegral)*

$$sieheForsterIII, S.54$$

.

Sei $f : \mathbb{R}^n \to \mathbb{R} \cup \{\pm\infty\}$ eine beliebige Funktion. Dann setzt man

$$\begin{aligned} \int^* f(x)dx &:= \inf\left\{ \int_{\mathbb{R}^n} \varphi(x)dx : \varphi \in \mathcal{H}^\uparrow(\mathbb{R}^n), \varphi \geq f \right\} , \\ \int_* f(x)dx &:= \sup\left\{ \int_{\mathbb{R}^n} \psi(x)dx : \psi \in \mathcal{H}^\downarrow(\mathbb{R}^n), \psi \leq f \right\} . \end{aligned}$$

Definition 9.2 *Eine Funktion $f : \mathbb{R}^n \to \mathbb{R} \cup \{\pm\infty\}$ heißt Lebesgue-integrierbar, falls*

$$-\infty < \int_* f(x)dx = \int^* f(x)dx < \infty .$$

Der gemeinsame Wert des Ober- und Unterintegrals heißt dann das Lebesgue-Integral von f und wird mit $\int f(x)dx$ bezeichnet.

Bemerkung 9.1 *Jede Funktion $f \in \mathcal{C}_c(\mathbb{R}^n)$ ist Lebesgue-integrierbar. Allgemeiner gilt: Eine Funktion $f \in \mathcal{H}^\uparrow(\mathbb{R}^n)$ bzw. $g \in \mathcal{H}^\downarrow(\mathbb{R}^n)$ ist genau dann Lebesgue-integrierbar, falls*

$$\int f(x)dx < \infty \ \ \textit{bzw.} \ \ \int g(x)dx > -\infty .$$

Sei $f : \mathbb{R}^n \longrightarrow \mathbb{R} \cup \{\pm\infty\}$ eine Funktion. Dann werden Funktionen $f_+, f_- : \mathbb{R}^n \longrightarrow \mathbb{R}_+ \cup \{\infty\}$ definiert durch

$$f_+(x) := \begin{cases} f(x), & \text{falls } f(x) \geq 0, \\ 0 & \text{sonst,} \end{cases}$$

$$f_-(x) := \begin{cases} -f(x), & \text{falls } f(x) \leq 0, \\ 0 & \text{sonst.} \end{cases}$$

Es gilt $f = f_+ - f_-$ und $|f| = f_+ + f_-$.

Satz 9.1 *Sei $f : \mathbb{R}^n \longrightarrow \mathbb{R} \cup \{\pm\infty\}$ eine Funktion.*

a) f ist genau dann integrierbar, wenn f_+ und f_- integrierbar sind.

b) Ist f integrierbar, so ist auch $|f|$ integrierbar.

Beweis: Forster, S. 60

Definition 9.3 *Wir bezeichnen mit $\mathcal{L}_1(\mathbb{R}^n)$ die Menge aller Lebesgue-integrierbaren Funktionen $f : \mathbb{R}^n \longrightarrow \mathbb{R}$. (Die Werte $\pm\infty$ sind hier nicht zugelassen.)*

Satz 9.2 *Seien $f, g \in \mathcal{L}_1(\mathbb{R}^n)$ und $\lambda \in \mathbb{R}$. Dann sind auch $\lambda f, f+g \in \mathcal{L}_1(\mathbb{R}^n)$ und es gilt:*

a) $\int (f(x)+g(x))dx = \int f(x)dx + \int g(x)dx$.

b) $\int \lambda f(x)dx = \lambda \int f(x)dx$.

c) Falls $f \leq g$, folgt $\int f(x)dx \leq \int g(x)dx$.

Beweis: Forster, S. 60

Bemerkung 9.2 *Satz 6.7 bedeutet, daß $\mathcal{L}_1(\mathbb{R}^n)$ ein Vektorraum ist und das Integral ein lineares monotones Funktional auf diesem Vektorraum darstellt.*

Satz 9.3 *a) Seien $f, g \in \mathcal{L}_1(\mathbb{R}^n)$. Die Funktion g sei beschränkt, d.h. es existiere eine Konstante $M \in \mathbb{R}_+$ mit $|g(x)| \leq M$ für alle $x \in \mathbb{R}^n$. Dann gilt $fg \in \mathcal{L}_1(\mathbb{R}^n)$.*

b) Sei $f \in \mathcal{L}_1(\mathbb{R}^n)$ eine beschränkte integrierbare Funktion und $p \geq 1$ eine reelle Zahl. Dann ist auch $|f|^p \in \mathcal{L}_1(\mathbb{R}^n)$.

Beweis: Forster, S. 61

Satz 9.4 *Seien $f_1 \in \mathcal{L}_1(\mathbb{R}^n)$ und $f_2 \in \mathcal{L}_1(\mathbb{R}^m)$. Dann gehört die durch*

$$(f_1 \otimes f_2)(x,y) := f_1(x)f_2(y)$$

für alle $x \in \mathbb{R}^n$, $y \in \mathbb{R}^m$ definierte Funktion $f_1 \otimes f_2 : \mathbb{R}^{n+m} \to \mathbb{R}$ zu $\mathcal{L}_1(\mathbb{R}^{n+m})$ und es gilt

$$\int\limits_{\mathbb{R}^{n+m}} f_1(x)f_2(y)d^n x d^m y = \int\limits_{\mathbb{R}^n} f_1(x)d^n x \cdot \int\limits_{\mathbb{R}^m} f_2(y)d^m y \ .$$

Satz 9.5 *Sei $f : \mathbb{R}^n \to \mathbb{R} \cup \{\pm\infty\}$ integrierbar und seien $A \in GL(n, \mathbb{R})$, $b \in \mathbb{R}^n$. Dann ist auch die Funktion $x \mapsto f(Ax+b)$ integrierbar und es gilt*

$$\int\limits_{\mathbb{R}^n} f(Ax+b)dx = \frac{1}{|\det A|} \int\limits_{\mathbb{R}^n} f(x)dx \ .$$

2. Meßbare Mengen

Definition 9.4 *Eine Teilmenge $M \subset \mathbb{R}^n$ heißt meßbar (oder integrierbar), falls ihre charakteristische Funktion χ_M integrierbar ist. In diesem Fall ist das Volumen (oder Lebesgue-Maß) von M definiert als*

$$\mathit{Vol}\,(M) := \int_{\mathbb{R}^n} \chi_M(x) d^n x \; .$$

Beispiel 9.1 *Jede kompakte Menge $K \subset \mathbb{R}^n$ ist integrierbar.*

Satz 9.6 *Seien $A, B \subset \mathbb{R}^n$ integrierbare Mengen. Dann sind auch die Mengen $A \cap B$, $A \cup B$ und $A \backslash B$ integrierbar und es gilt*

$$\begin{aligned} \mathit{Vol}\,(A \cup B) &= \mathit{Vol}\,(A) + \mathit{Vol}\,(B) - \mathit{Vol}\,(A \cap B) \; , \\ \mathit{Vol}\,(A \backslash B) &= \mathit{Vol}\,(A) - \mathit{Vol}\,(A \cap B) \; . \end{aligned}$$

Beispiel 9.2 *Jede beschränkte offene Menge $U \subset \mathbb{R}^n$ ist integrierbar. Denn die abgeschlossene Hülle $\bar{U}$ und der Rand ∂U sind kompakt, also integrierbar. Deshalb ist auch $U = \bar{U} \backslash \partial U$ integrierbar.*

Integration über Teilmengen

Definition 9.5 *Sei $M \subset \mathbb{R}^n$ eine beliebige Teilmenge. Eine Funktion $f : M \to \mathbb{R} \cup \{\pm\infty\}$ heißt integrierbar über M, falls die trivial fortgesetzte Funktion $\tilde{f} : \mathbb{R}^n \to \mathbb{R} \cup \{\pm\infty\}$,*

$$\tilde{f}(x) := \begin{cases} f(x) & \text{für} \quad x \in M \; , \\ 0 & \text{für} \quad x \in \mathbb{R}^n \backslash M \; , \end{cases}$$

über $\mathbb{R}^n$ integrierbar ist. Man setzt dann

$$\int_M f(x) dx := \int_{\mathbb{R}^n} \tilde{f}(x) dx \; .$$

Definition 9.6 *Eine Teilmenge $M \subset \mathbb{R}^n$ heißt Nullmenge, wenn sie integrierbar ist und das Lebesgue-Maß (Volumen) Null hat.*

Bemerkung 9.3 *Da das Unterintegral der charakteristischen Funktion einer beliebigen Menge stets nicht-negativ ist, ist M genau dann Nullmenge, wenn*

$$\int^* \chi_M(x) dx = 0 \; .$$

Satz 9.7 *a) Sei $M \subset \mathbb{R}^n$ eine Nullmenge und $N \subset M$. Dann ist auch N eine Nullmenge.*

b) Sei $M_k \subset \mathbb{R}^n$, $k \in \mathbb{N}$, eine abzählbare Familie von Nullmengen. Dann ist auch die Vereinigung

$$M := \bigcup_{k=0}^{\infty} M_k$$

eine Nullmenge.

Satz 9.8 *Sei $A \subset \mathbb{R}^n$ eine Nullmenge. Dann gibt es zu jedem $\epsilon > 0$ eine integrierbare offene Menge $U \subset \mathbb{R}^n$ mit*

$$A \subset U \quad \textit{und} \quad \mathit{Vol}\,(U) < \epsilon\,.$$

Satz 9.9 *Eine Teilmenge $A \subset \mathbb{R}^n$ ist genau dann eine Nullmenge, wenn es zu jedem $\epsilon > 0$ eine abzählbare Familie $(Q_i)_{i\in I}$ von Quadern (oder Würfeln) $Q_i \subset \mathbb{R}^n$ gibt mit*

$$A \subset \bigcup_i Q_i \quad \textit{und} \quad \sum_i \mathit{Vol}\,(Q_i) < \epsilon\,.$$

Satz 9.10 *Sei $U \subset \mathbb{R}^n$ offen und $F : U \to \mathbb{R}^n$ eine stetig differenzierbare Abbildung. Dann ist für jede Nullmenge $A \subset U$ das Bild $F(A)$ eine Nullmenge.*

Beispiele von Nullmengen

1. Jeder Punkt im $\mathbb{R}^n (n \geq 1)$ ist eine Nullmenge. Also ist auch jede abzählbare Punktmenge eine Nullmenge.

2. Jede Hyperebene $H \subset \mathbb{R}^n$ ist eine Nullmenge.

3. Der Rand jedes Polyeders ist eine Nullmenge.

4. Sei $A \subset \mathbb{R}^{n-1}$ eine kompakte Menge und $f : A \longrightarrow \mathbb{R}$ eine stetige Funktion. Dann ist der Graph Γ von f ,

$$\Gamma := \{(x, y) \in \mathbb{R}^{n-1} \times\ \mathbb{R} : x \in A\,,\ \ y = f(x)\}$$

eine Nullmenge in $\mathbb{R}^n$.

Definition 9.7 *Gegeben seien zwei Funktionen $f, g : \mathbb{R}^n \to \mathbb{R} \cup \{\pm\infty\}$. Wir nennen f und g (Lebesgue-)fast überall gleich, wenn die Menge*

$$\{x \in \mathbb{R}^n : f(x) \neq g(x)\}$$

eine Nullmenge ist.

Bemerkung 9.4 *Sei $f = g$ fast überall. Falls im Punkt $a \in \mathbb{R}^n$ beide Funktionen endlich und stetig sind, gilt $f(a) = g(a)$. Denn in jeder Umgebung U von a gibt es Punkt $x \in U$ mit $f(x) = g(x)$.*

Satz 9.11 *Seien $f, g : \mathbb{R}^n \to \mathbb{R} \cup \{\pm\infty\}$ zwei Funktionen, so daß*

$$f = g \quad \textit{fast überall}\,.$$

Ist f integrierbar, so ist auch g integrierbar und es gilt

$$\int f(x)dx = \int g(x)dx\,.$$

Sei $A \subset \mathbb{R}^n$ eine Menge und $N \subset A$ eine Nullmenge. Gilt eine Aussage a für $x \in A \backslash N$, dann sagt man, daß die Aussage fast überall gilt und schreibt:

$$a \quad \text{fast überall}\,.$$

“fast überall” wird auch als “presque partout” oder “p.p.” oder “almost everywhere” oder “a.e.” geschrieben.

Satz 9.12 *(Fubini):*
Sei $f : \mathbb{R}^k \times \mathbb{R}^m \to \mathbb{R} \cup \{\pm\infty\}$ eine integrierbare Funktion. Dann gibt es eine Nullmenge $N \subset \mathbb{R}^m$, so daß für jedes fest $y \in \mathbb{R}^m \backslash N$ die Funktion

$$\begin{array}{rcl} \mathbb{R}^k & \longrightarrow & \mathbb{R} \cup \{\pm\infty\} \\ x & \longmapsto & f(x,y) \end{array}$$

integrierbar ist. Setzt man

$$F(y) := \int\limits_{\mathbb{R}^k} f(x,y)d^kx \quad \textit{für} \quad y \in \mathbb{R}^m \backslash N$$

und definiert $F(y)$ für $y \in N$ beliebig, so ist die Funktion $F : \mathbb{R}^m \to \mathbb{R} \cup \{\pm\infty\}$ integrierbar und es gilt

$$\int\limits_{\mathbb{R}^{k+m}} f(x,y)d^kxd^my = \int\limits_{\mathbb{R}^m} F(y)d^my\,.$$

3. Konvergenzsätze

Satz 9.13 *(Satz von der monotonen Konvergenz von B. Levi).*
Sei $f_k : \mathbb{R}^n \to \mathbb{R} \cup \{\pm\infty\}$, $k \in \mathbb{N}$, eine Folge integrierbarer Funktionen mit $f_k \leq f_{k+1}$ für alle $k \in \mathbb{N}$. Falls

$$\lim_{k\to\infty} \int f_k(x)dx =: M < \infty ,$$

so ist die Funktion

$$f := \lim_{k\to\infty} f_k$$

integrierbar und es gilt

$$\int f(x)dx = \lim_{k\to\infty} \int f_k(x)dx .$$

Satz 9.14 *(Satz von der majorisierten Konvergenz von H. Lebesgue).*
Sei $f_k \in \mathcal{L}_1(\mathbb{R}^n)$, $k \in \mathbb{N}$, eine Folge integrierbarer Funktionen, die fast überall auf $\mathbb{R}^n$ punktweise gegen eine Funktion $f : \mathbb{R}^n \to \mathbb{R}$ konvergiere. Es gebe eine Funktion $F : \mathbb{R}^n \to \mathbb{R}_+ \cup \{\infty\}$ mit $\|F\|_{L_1} < \infty$, so daß

$$|f_k| \leq F \quad \textit{für alle} \quad k \in \mathbb{N} .$$

Dann ist auch f integrierbar und

$$\int f(x)dx = \lim_{k\to\infty} \int f_k(x)dx .$$

9.3 Topologische Räume

In diesem Absatz besprechen wir einige Begriffe bzgl. topologischer Räume.

<u>Topologische Räume</u>

Sei X eine Menge. Ein System τ von Untermengen von X,

$$\tau \subset P^X$$

ist eine Topologie auf X, falls:

1. $\emptyset \in \tau$
2. $X \in \tau$
3. Seien $\alpha_1, \ldots , \alpha_m \in \tau$. Dann gilt

$$\bigcap_{i=1}^{n} \alpha_i \in \tau .$$

D.h. der Durchschnitt endlich vieler Mengen aus τ gehört zu τ.

4. Sei $\alpha_i \in \tau$, $i \in I$. Dann gilt

$$\bigcup_{i \in I} \alpha_i \in \tau \ .$$

   D.h. die Vereinigung beliebig vieler Mengen aus τ liegt in τ.

Die Elemente $\alpha \in \tau$ heißen *offene Mengen* des topologischen Raumes (X, τ).

Eine Teilmenge $M \subset X$ heißt *Umgebung* eines Punktes x des topologischen Raumes (X, τ), falls

1. $x \in M$
2. Es existiert $U \in \tau$ mit

$$x \in U \subset M \ .$$

Sei X ein linearer Vektorraum. Sei τ eine Topologie auf X. (X, τ) heißt *topologischer Vektorraum*, falls Addition und Multiplikation, d.h. die Abbildungen

$$\begin{aligned} (x,y) &\longmapsto x+y \ , \\ (\lambda, x) &\longmapsto \lambda x \ , \end{aligned}$$

stetig sind.

Eine Menge $\mathcal{B}$ von Umgebungen von 0 heißt *Nullumgebungsbasis*, wenn jede Umgebung U von 0 ein $V \in \mathcal{B}$ umfaßt.

Ein topologischer Vektorraum heißt *lokal konvex*, falls jede Umgebung eines Punktes eine konvexe Umgebung des Punktes umfaßt.

Satz 9.15 *Ein lokalkonvexer Vektorraum X besitzt eine Nullumgebungsbasis mit:*

1. *Für jedes $U, V \in U$ gibt es ein $W \in U$ mit*

$$W \subset U \cap V \ .$$

2. *Für jedes $U \in U$ und $\alpha \neq 0$ gilt $\alpha U \in U$.*
3. *Jedes $U \in U$ ist balanciert, konvex und absorbierend, d.h.*

   (a) $u, v \in U$, $\lambda \in [0,1] \Longrightarrow \lambda u + (1-\lambda) v \in U$. (konvex)

(b) *Für jedes* $u \in X$ *existiert* $\alpha \in \mathbb{R}$ *mit* $\alpha^{-1}u \in U$. *(absorbierend)*

(c) $u \in U$, $|\alpha| \leq 1 \Longrightarrow \alpha U \in U$. *(balanciert)*

Umgekehrt, jede Menge U mit den Eigenschaften 1 - 3 erzeugt eine Topologie τ, so daß (X, τ) lokalkonvex ist.
Sei

$$\begin{aligned} \mathcal{V} &:= \{V \subset X : U \subset V \quad \text{für mindestens ein } U \in \tau\} \\ \mathcal{V}_x &= \{x + V : V \in \mathcal{V}\} . \end{aligned}$$

Dann ist $\mathcal{V}_x$ eine Umgebungsbasis für x.

Beweis: Robertson + Robertson, S. 10.

Sei X ein Vektorraum. Eine Abbildung $p : X \longrightarrow \mathbb{R}_+$ heißt *Halbnorm*, falls:

1. $p(x) + p(y) \leq p(x) + p(y)$
2. $p(\alpha x) = |\alpha| p(x)$.

Satz 9.16 *Sei* p *eine Halbnorm auf* X *und* $\epsilon \in \mathbb{R}$, $\epsilon > 0$. *Sei*

$$M := \{x \in X : p(x) \leq \epsilon\} .$$

Dann gilt:

1. $O \in M$.
2. *M ist konvex.*
3. *M ist balanciert, d.h. sei $x \in M$, $|\alpha| \leq 1$, dann folgt*
$$\alpha x \in M .$$
4. *M ist absorbierend, d.h. für jedes $x \in X$ existiert $\alpha \in \mathbb{R}$ mit*
$$\alpha^{-1} x \in M .$$
5. $p(x) = \inf\{\alpha > 0 : x \in \alpha M\}$.

Beweis: Robertson + Robertson, S. 13.

Sei $\{p_\gamma(x) : \gamma \in \Gamma\}$ eine Familie von Halbnormen auf einem Vektorraum X der der Eigenschaft, daß für jedes $x_0 \in X$ eine Halbnorm p_γ existiert mit

$$p_\gamma(x_0) \neq 0 \ .$$

Sei

$$U = U(\epsilon_1, \dots, \epsilon_n \ ; \ \gamma_1, \dots, \gamma_n) = \{x \in X : p_{\gamma_k}(x) \leq \epsilon_k \ , \ 1 \leq k \leq n\} \ .$$

U ist konvex, balanciert und absorbierend. Sei U_0 die Menge aller U.

Behauptung: U_0 ist eine Nullumgebungsbasis für eine Topologie τ. (X, τ) ist ein lokalkonvexer topologischer Raum.

Beweis: Siehe Yosida, S. 25, Robertson, S. 15.

Sei X ein Vektorraum. Sei $\{X_\alpha\}$ eine Familie von Unterräumen, $X_\alpha \subset X$, mit $X = \cup X_\alpha$. Jedes X_α sei ein lokalkonvexer topologischer Vektorraum, so daß - wenn $X_\alpha \subset X_\beta$ - gilt: Die Topologie von X_α ist mit der relativen Topologie von X_α bzgl. X_β identisch. U_0 sei die Menge von allen Mengen $U_0 \subset X$ mit:

1. U_0 ist konvex und balanciert.
2. $U_0 \cap X_\alpha$ ist eine Umgebung von 0 in X_α.

Behauptung: U_0 ist eine Nullumgebungsbasis für eine Topologie τ. Mit dieser Topologie τ ist (X, τ) ein lokalkonvexer topologischer Vektorraum. τ heißt die *induktive Topologie* von X bzgl. $\{X_\alpha\}$.

Beweis: Siehe Robertson + Robertson, S. 79.

9.4 Sobolev-Räume: Definition durch Vervollständigung

Sei Ω eine offene Teilmenge des $\mathbb{R}^n$. Sei

$C^\infty(\Omega) :=$ Menge aller unendlich oft differenzierbaren reellen Funktionen auf Ω.

Satz 9.17 *Sei X ein metrischer Raum. Dann gibt es einen Fréchet-Raum Y und einen Unterraum $\tilde{X} \subset Y$ mit:*

1. $\tilde{X}$ ist dicht in Y.

2. X und $\tilde{X}$ sind isometrisch und isomorph.

Der Raum $\tilde{X}$ heißt Vervollständigung von X und ist bis auf Isometrie eindeutig.

Beweis: Der Beweis basiert auf Cauchy-Folgen, wie bei der Vervollständigung von Q zu $\mathbb{R}$.

Definition 9.8 *Sei Ω eine offene Teilmenge des $\mathbb{R}_n$, $n \in \mathbb{N}$, $p \in \mathbb{R}$, $m \geq 0$, $p \geq 1$. Sei $L^p(\Omega)$ der Lebesguesche Raum. Die Elemente dieses Raumes sind Klassen äquivalenter Funktionen $x(t)$, die auf Ω Lebesgue-meßbar sind und für die $|x(t)|^p$ Lebesgue-summierbar ist:*

$$\begin{aligned}
\|x\|_p &= \left[\int_\Omega |x(t)|^p dt\right]^{1/p} , \quad 1 \leq p < \infty , \\
\|x\|_\infty &= \textit{vrai} \max_{t\in\Omega} |x(t)| = \textit{ess.} \sup |x(t)| \\
&= \inf\{K : |x(t)| \leq K \quad \textit{fast überall}\} . \\
\|u\|_{m,p} &= \left[\sum_{0\leq|\alpha|\leq m} (\|D^\alpha u\|_p)^p\right]^{1/p} , \quad 1 \leq p < \infty , \\
\|u\|_{m,\infty} &= \max_{|\alpha|\leq m} \|D^\alpha u\|_\infty . \\
H^{m,p}(\Omega) &:= \textit{Vervollständigung von } \{u \in C^\infty(\Omega) : \|u\|_{m,p} < \infty\} \\
H_0^{m,p}(\Omega) &:= \textit{Vervollständigung von } \{u \in C_0^\infty(\Omega) : \|u\|_{m,p} < \infty\} \quad \textit{bezüglich } \|\cdot\|_{m,p} .
\end{aligned}$$

Diese Definitionen von $H^{m,p}(\Omega)$ und $H_0^{m,p}(\Omega)$ sind sehr einfach und oft sehr nützlich. Man kann sie z.B. benutzen, um die Definitionsbereiche der Bilinearform a und des linearen Funktionals b zu erweitern:

$$a(u,v) := \lim_{\substack{\varphi_n \to u \\ \Psi_n \to v \\ \varphi_n, \Psi_n \in C^\infty(\Omega)}} a(\varphi_n, \Psi_n) .$$

Es läßt sich leicht feststellen, daß $\{a(\varphi_n, \Psi_n)\}$ eine Cauchy-Folge ist, wenn $\{\varphi_n\}$ und $\{\Psi_n\}$ Cauchy-Folgen sind.

9.5 Sobolev-Räume: Definition mit Hilfe schwacher Ableitungen

Obwohl die Definition von $H^{m,p}(\Omega)$ und $H_0^{m,p}(\Omega)$ mit Hilfe der Vervollständigung sehr schlicht ist, hat sie den Nachteil, daß die funktionstheoretischen Eigenschaften der

Elemente von $H^{m,p}(\Omega)$ nicht geklärt sind. In diesem Absatz wird deshalb ein ganz anderer Weg eingeschlagen.

Sei Ω eine offene Teilmenge von $\mathbb{R}^n$. Sei

$$\begin{aligned} X &:= C_0^\infty(\Omega) \\ &= \text{die Menge aller unendlich oft differenzierbaren Funktionen auf} \\ &\quad\ \Omega \text{ mit kompaktem Träger} \end{aligned}$$

Sei $K \subset \Omega$ kompakt,

$$\mathcal{D}_K(\Omega) = \{\phi \in X : \operatorname{supp}(\phi) \subset K\}\ .$$

Sei

$$p_{K,m}(\phi) := \sup_{\substack{|\alpha| \leq m \\ x \in K}} |D^\alpha \phi|\ ,$$

für $\phi \in \mathcal{D}_K(\Omega)$, $m \in \mathbb{N}_0$, $m \geq 0$. $p_{K,m}$ ist eine Seminorm auf $\mathcal{D}_K(\Omega)$. Seien $\epsilon_1, \ldots, \epsilon_n$ positive Zahlen,

$$U(\epsilon_1, \ldots, \epsilon_n\ ;\ m_1, \ldots, m_n\ ;\ K) = \{\phi \in \mathcal{D}_K(\Omega) : p_{K,m_j}(\phi) \leq \epsilon_j\ ,\ \ 1 \leq j \leq n\}\ .$$

Mit dieser Nullumgebungsbasis ist $X_K = \mathcal{D}_K(\Omega)$ ein lokalkonvexer topologischer Vektorraum.

Definition 9.9 *Mit der induktiven Topologie τ bzgl. $\{X_K\}$ ist (X, τ) ein lokalkonvexer topologischer Raum, der mit $\mathcal{D}(\Omega)$ bezeichnet wird.*

Bemerkung 9.5 *Siehe Yosida, S. 28, für weitere Eigenschaften von $\mathcal{D}(\Omega)$.*

Der Raum $\mathcal{D}'(\Omega)$, der duale Raum zu $\mathcal{D}(\Omega)$, besteht aus allen stetigen linearen Funktionalen auf $\mathcal{D}(\Omega)$. Ein Element $T \in \mathcal{D}'(\Omega)$ heißt *Distribution* oder *Verallgemeinerte Funktion.* $\varphi \in \mathcal{D}(\Omega)$ heißt *Testfunktion.*
Seien nun $S, T, T_n \in \mathcal{D}'(\Omega)$ und $c \in \mathbb{R}$. Es gilt (siehe z.B. Adams, S. 19):

1. $(S + T)(\varphi) = S(\varphi) + T(\varphi), \forall \varphi \in \mathcal{D}(\Omega)$.
2. $(cT)(\varphi) = c(T(\varphi))\ ,\ \forall \varphi \in \mathcal{D}(\Omega)$.
3. $T_n \longrightarrow T$ in $\mathcal{D}'(\Omega)$ genau dann, wenn $T_n\varphi \longrightarrow T\varphi$ für alle $\varphi \in \mathcal{D}(\Omega)$.

Die Topologie von $\mathcal{D}'(\Omega)$ ist die schwach*-Topologie, d.h. die schwächste Topologie, womit alle Funktionale L_φ , $\varphi \in \mathcal{D}(\Omega)$,

$$\begin{aligned} L_\varphi &: \mathcal{D}'(\Omega) \longrightarrow \mathbb{R} , \\ L_\varphi(T) &= < T, \varphi >:= T(\varphi) \end{aligned}$$

stetig sind. Diese Topologie wird durch die Seminorm

$$p(T) = p(T; \varphi_1, \dots, \varphi_m) = \sup_{1 \le k \le m} |T(\varphi_k)|$$

erzeugt.

Definition 9.10 *Sei $T \in \mathcal{D}'(\Omega)$, α ein Multiindex. Die Ableitung $D^\alpha T$ ist die Distribution*

$$(D^\alpha T)(\varphi) = (-1)^{|\alpha|} T(D^\alpha \varphi) .$$

Sei $u \in L_1^{\text{loc}}(\Omega)$, d.h. $u|_A \in L_1(A)$ für jede meßbare Teilmenge $A \subset \Omega$, wofür $\bar{A} \subset \Omega$ und $\bar{A}$ kompakt ist. Sei

$$T_u \varphi := \int_\Omega u(x)\varphi(x)dx .$$

T_u ist ein lineares Funktional. Weiter ist T_u stetig. D.h. T_u ist eine Distribution, $T_u \in \mathcal{D}'(\Omega)$.

Definition 9.11 *Sei $u \in L_1^{loc}(\Omega)$. $v \in L_1^{loc}(\Omega)$ heißt die schwache Ableitung $D^\alpha u$ von u, falls*

$$T_v = D^\alpha(T_u) ,$$

d.h.

$$\int_\Omega v(x)\varphi(x)dx = (-1)^{|\alpha|} \int u(x) D^\alpha \varphi(x)dx \quad \textit{für alle} \ \ \varphi \in \mathcal{D}(\Omega) .$$

Beispiel 9.3 *Sei H die Heaviside-Funktion*

$$H(x) = \begin{cases} 1 & , \quad x \ge 0 \\ 0 & , \quad x < 0 \end{cases}$$

$$(D(T_H))(\varphi) = \varphi(0) = \delta_0 \varphi ,$$

wo δ_0 die Diracsche Funktion ist.

$$\delta_0 \varphi := \varphi(0) .$$

In diesem Beispiel hat die Funktion H keine verallgemeinerte Ableitung.

Bemerkung 9.6 *Falls die verallgemeinerte Ableitung $D^\alpha u$ existiert, dann ist sie bis auf Nullmengen eindeutig bestimmt.*

Definition 9.12 *Sei $m \in \mathbb{N}_0$, $p \in \mathbb{R}$, $p \geq 1$, $\Omega \subset \mathbb{R}^n$. $W^{m,p}(\Omega)$ sei die Menge aller Funktionen u, die schwache Ableitungen $D^\alpha u \in L^p(\Omega)$ besitzen. Man setzt*

$$\|u\|_{m,p} = \left[\sum_{|\alpha| \leq m} (\|D^\alpha u\|_p)^p \right]^{1/p} .$$

Satz 9.18 $W^{m,p}(\Omega)$ *ist ein Banachraum.*

Beweis: (siehe Yosida, S. 55).

Sei $\{u_k\}$ eine Cauchy-Folge in $W^{m,p}(\Omega)$. D.h., für jedes $\epsilon > 0$ gibt es N mit

$$\|u_n - u_m\|_{m,p} \leq \epsilon \quad \text{für} \quad m, n \geq N .$$

Es folgt, daß die Folgen von schwachen Ableitungen $\{D^\alpha u_k\}$ auch Cauchy-Folgen sind, so daß Elemente $f^{(\alpha)} \in L_p(\Omega)$ existieren mit

$$D^\alpha u_k \longrightarrow f^{(\alpha)} \quad \text{in} \quad L_p(\Omega) .$$

Es mußdann gezeigt werden, daß $f^{(\alpha)} = D^\alpha f^{(0)}$ gilt.

9.6 Das Verhältnis zwischen $W^{m,p}(\Omega)$ und $H^{m,p}(\Omega)$

Sei

$$X = C^\infty(\Omega) \cap H^{m,p}(\Omega) .$$

X ist dicht in $H^{m,p}(\Omega)$. Jedes Element $x \in X$ liegt auch in $W^{m,p}(\Omega)$. Es folgt:

$$H^{m,p}(\Omega) \subset W^{m,p}(\Omega) .$$

Lange Zeit blieb die Frage offen, unter welchen Bedingungen $H^{m,p}(\Omega) = W^{m,p}(\Omega)$. 1964 zeigten Meyers und Serrin in einer sehr schönen Arbeit, daß folgendes gilt:

Satz 9.19 *Sei $1 \leq p < \infty$, $m \in \mathbb{N}_0$. Dann gilt:*

$$H^{m,p}(\Omega) = W^{m,p}(\Omega) .$$

Beweis: Siehe Adams, S. 52.

Der Satz ist falsch für $p = \infty$.

Beispiel 9.4 *Sei* $n = 1$, $p = 1$, $\Omega = (-1, +1)$, $u(x) := |x|$.

Die schwache Ableitung Du *existiert, da für jedes* $\varphi \in C_0^\infty(-1, +1)$,

$$
\begin{aligned}
(-1)\int_{-1}^{+1} u\, \varphi'(x)dx &= \int_{-1}^{0} x\, \varphi'(x)dx - \int_{0}^{1} x\, \varphi'(x)dx \\
&= [x\, \varphi(x)]_{-1}^{0} - \int_{-1}^{0} \varphi(x)dx - [x\, \varphi(x)]_0^1 + \int_0^1 \varphi(x)dx \\
&= \int_{-1}^{+1} v(x)\varphi(x)dx
\end{aligned}
$$

mit

$$
v(x) = \begin{cases} -1 & , \quad x \leq 0 \\ +1 & , \quad x > 0 \end{cases}
$$

D.h. $v = Du$, $u \in W^{1,\infty}(-1, +1)$ *und*

$$
\|u\|_{1,\infty} = \max_{0 \leq \alpha \leq 1} \|D^\alpha u\|_\infty = 1 \, .
$$

Ist $u \in H^{1,\infty}(-1, +1)$*? Wenn ja, dann existiert eine Cauchy- Folge* $\{\varphi_n\}$ *mit*

$$
\varphi_n \in C^\infty(-1, +1)
$$

und

$$
\begin{aligned}
\|\varphi_n - u\|_\infty &\longrightarrow 0 \, , \\
\|\varphi_n' - v\|_\infty &\longrightarrow 0 \, ,
\end{aligned}
$$

wo $v \in L^1$. *Es läßt sich dann zeigen, daß* $v = Du$.

Sei aber $\epsilon < 1$. *In jedem Intervall* $(-\delta, +\delta)$ *ist*

$$
\begin{aligned}
\max\, Du &= 1 \, , \\
\min\, Du &= 1 \, .
\end{aligned}
$$

Es folgt, daß in so einem Intervall

$$
\|\varphi_n' - Du\|_\infty > 1 - \epsilon \, .
$$

9.7 Zusammenhang mit absoluten stetigen Funktionen

Definition 9.13 *Eine auf dem Segment $[a,b]$ definierte endliche Funktion $f(x)$ heißt absolut stetig, wenn zu jedem $\epsilon > 0$ ein $\delta > 0$ existiert, so daß für ein beliebiges endliches System von paarweise durchschnittsfremden Intervallen $(a_1,b_1),\dots,(a_n,b_n)$ mit der Gesamtlänge*

$$\sum_{k=1}^{n}(b_k - a_k) < \delta$$

gilt

$$\left|\sum_{k=1}^{n}\{f(b_k) - f(a_k)\}\right| < \epsilon\ .$$

Satz 9.20 *Jede absolut stetige Funktion ist unbestimmtes Integral ihrer Ableitung.*

Beweis: Siehe Natanson, S. 257.

Satz 9.21 *(siehe Smirnov, S. 3326.)*
Sei $\Omega := \{x : 0 < x_i < 1\ ,\ 1 \le i \le n\}$. Sei $u \in W^{1,p}(\Omega)$ und $p > 1$.
Dann gibt es eine Funktion φ, die auf $\bar{\Omega}$ definiert ist, so daß

1. *φ ist äquivalent zu u auf Ω, d.h. $\varphi(x) = u(x)$ f.ü.*

2. *φ ist absolut stetig bzgl. x_1 für $0 \le x_1 \le 1$ für fast alle $x_2,\dots,x_n$.*

3. *Für alle $x_i \in [0,1]$ gilt:*

$$\varphi(x) = \varphi(0,x_2,\dots,x_n) + \int_0^{x_1} \frac{\partial\varphi(t,x_2,\dots,x_n)}{\partial t}dt\ .$$

Morrey, S. 66, untersucht weitere Zusammenhänge zwischen verallgemeinerten Ableitungen und absoluter Stetigkeit.

9.8 Die Sobolewschen Einbettungssätze

9.9 Der Spur Operator γ

9.10 Die Poincarésche Ungleichung und ihre Anwendungen

Sei

$$|u|_{m,p} = \left[\sum_{|\alpha|=m} (\|D^\alpha u\|_p)^p \right]^{1/p} .$$

Auf $H^{m,p}(\Omega)$ ist $|u|_{m,p}$ eine Halbnorm, da $|1|_{m,p} = 0$. Für $H_0^{m,p}(\Omega)$ ist $|u|_{m,p}$ aber eine Norm, die zu $\|u\|_{m,p}$ äquivalent ist, wie der folgende Satz zeigt:

Satz 9.22 *Sei $\Omega \subset \mathbb{R}^n$, $d > 0$. Ω liege zwischen den Ebenen $x_n = 0$ und $x_n = d$. Dann gibt es eine Konstante K mit*

$$|u|_{m,p} \le \|u\|_{m,p} \le K|u|_{m,p} \quad \textit{für alle } u \in H_0^{m,p} .$$

<u>Beweis:</u> (Adams, S. 158)

Sei $\varphi \in C_0^\infty(\Omega)$. Man schreibe $x = (x', x_n)$ mit $x' \in \mathbb{R}^{n-1}$. Dann gilt:

$$\varphi(x) = \int_0^{x_n} \frac{d}{dt}\varphi(x', t)dt . \qquad (*)$$

Die Ungleichung von Hölder besagt:

$$\int_\Omega |u(x)v(x)|dx \le \|u\|_p \, \|v\|_p \quad \text{für } u \in L^p(\Omega) \, , \; v \in L^{p'}(\Omega) \, , \; \frac{1}{p} + \frac{1}{p'} = 1 .$$

Satz 9.23 *Sei*

$$1 < p < \infty \, , \; u \in L^p(\Omega) \, , \; v \in L^{p'}(\Omega) \, , \; \frac{1}{p} + \frac{1}{p'} = 1 .$$

Dann ist $uv \in L^1(\Omega)$ und

$$\int_\Omega |u(x)v(x)|dx \le \|u\|_p \cdot \|v\|_{p'} .$$

<u>Beweis:</u>

Sei

$$f(t) = \frac{t^p}{p} + \frac{1}{p'} - t .$$

Für $t \ge 0$ ist $f \ge 0$ mit $f(1) = 0$. Sei $t = ab^{-p'/p}$. Dann folgt:

$$ab \le \frac{a^p}{p} + \frac{b^{p'}}{p'} \qquad \text{(Young's Ungleichung)} .$$

Sei $\|u\|_p \, , \; \|v\|_{p'} > 0$. Man setze

$$a := \frac{|u(x)|}{\|u\|_p}\ ,\ b := \frac{|v(x)|}{\|v\|_{p'}}\ .$$

Mit $u = 1$ und $v = \frac{d}{dt}\varphi(x', t)$ folgt aus $(*)$:

$$\begin{aligned}
|\varphi(x)|^p &= \left[\int_0^{x_n} u(t)v(t)dt\right]^p \\
&\leq [\|u\|_{p'} \cdot \|v\|_p]^p \\
&= \left(\int_0^{x_n} 1\ dt\right)^{\frac{p}{p'}} \cdot \int_0^{x_n} \left|\frac{d}{dt}\varphi(x', t)\right|^p dt \\
&\leq x_n^{p-1} \cdot \int_0^d \left|\frac{d}{dt}\varphi(x', t)\right|^p dt\ .
\end{aligned}$$

Es folgt

$$\begin{aligned}
(\|\varphi\|_{0,p})^p &= \int_{\mathbb{R}^{n-1}} dx' \int_0^d |\varphi(x)|^p dx_n \\
&\leq \int_{\mathbb{R}^{n-1}} dx' \int_0^d x_n^{p-1} dx_n \int_0^d \left|\frac{d}{dt}\varphi(x', t)\right|^p dt \\
&\leq \frac{d^p}{p} \int_{\mathbb{R}^n} dx' \int_0^d \left|\frac{d}{dt}\varphi(x', t)\right|^p dt \\
&\leq \frac{d^p}{p} (|\varphi|_{1,p})^p
\end{aligned}$$

Zusammenfassend

$$|\varphi|_{1,p}^p \leq \|\varphi\|_{1,p}^p = \|\varphi\|_{0,p}^p + |\varphi|_{1,p}^p \leq \left(\frac{d^p}{p} + 1\right) |\varphi|_{1,p}^p\ . \qquad (**)$$

Die Ungleichung $(**)$ gilt auch, wenn φ durch ihre Ableitungen ersetzt wird. Der Satz folgt nun z.B. durch Induktion.

Bemerkung 9.7 *1. Damit $\|\varphi\|_{1,p} \leq const|\varphi|_{1,p}$ ist nur erforderlich, daß $\varphi = 0$ auf einem Teil Γ_1 von $\partial\Omega$ mit positivem Oberflächenmaß.*

2. Das Problem

$$\underset{\varphi \in H_0^1(\Omega)}{\mathit{Minimiere}} \quad \frac{|\varphi|_{0,2}}{|\varphi|_{1.2}}$$

führt zu einer Eigenwertaufgabe. Siehe Polya und Szego [1951].

3. *Ungleichungen Poincaréscher Art für Differenzen wurden schon in der Arbeit von Courant, Friedrichs und Lewy angewandt. Siehe auch das Buch von Epstein.*

Als direkte Folgerung aus der Poincaréschen Ungleichung erhält man:

Satz 9.24 *Sei $\Omega \subset \mathbb{R}^n$ beschränkt. Die Normen $\|\cdot\|_{1,p}$ und $|\cdot|_{1,p}$ sind äquivalente Normen in $H_0^{1,p}$.*

Satz 9.25 *Sei $\Omega \subset \mathbb{R}^n$ beschränkt. Sei*

$$\begin{aligned} a(u,v) &:= \int_\Omega div\, u \cdot div\, v \, dx \\ &= \int_\Omega \sum_{i=1}^n \frac{\partial u}{\partial x_i} \frac{\partial v}{\partial x_i} \, dx \end{aligned}$$

für $u, v \in H_0^1(\Omega)$.
a ist eine beschränkte V-elliptische Bilinearform.

Literatur

Adams, R.A.: Sobolev Spaces. New York: Academic Press, 1975.

Ciarlet, P.G.: The Finite Element Method for Elliptic Problems. North-Holland, 11978.

Gilbarg, D., **Trudinger**, N.S.: Elliptic Partial Differential Equations of Second Order. Springer 1977.

Hackbusch, W.: Theorie und Numerik elliptischer Differentialgleichungen. Teubner 1986.

Heuser, H.: Funktionalanalysis. Teubner 1986.

Kelley, J.L. **Namioka**, I.: Linear Topological Spaces. Springer 1963.

Morrey, C.B. Jr.: Multiple Integrals in the Calculus of Variations. Springer 1966.

Natanson, I.P.: Theorie der Funktionen einer reellen Veränderlichen.

Robertson, A.P., **Robertson**, W.J.: Topological Vector Spaces. Cambridge 1964.

Smirnow, W.I.: Lehrgang der Höheren Mathematik, Teil V, VEB Deutscher Verlag der Wissenschaften 1962.

Sobolev, S.L.: Applications of Functional Analysis in Mathematical Physics. American Mathematical Society 1963.

Schwarz, H.R.: Methode der finiten Elemente. Teubner 1980.

Strang, G.J.: Fix: An Analysis of the Finite Element Method. Prentice-Hall 1973.

Yosida, K.: Functional Analysis, 2nd edition. Springer 1968.

Zienkiewicz, O.C.: The Finite Element Method in Engineering Science. Second edition, McGraw-Hill 1971. (Es gibt eine neuere Ausgabe.)

Kapitel 10

Finite Elemente

10.1 Einführung

10.2 Das Galerkin-Verfahren: Fehlerabschätzungen

Nach diesen langen, aber durchaus spannenden Vorbereitungen können wir uns wieder der Lösung von Randwertaufgaben widmen.

Satz 10.1 *Sei X ein reeller Hilbertraum. Sei a und b eine Bilinearform*

$$a : X \times X \longrightarrow \mathbb{R}$$

bzw. ein lineares Funktional

$$b : X \longrightarrow \mathbb{R} \, .$$

Sei a beschränkt und X-elliptisch:

$$\begin{aligned} a(u,v) &\leq M\|u\| \cdot \|v\| \, , \\ a(u,u) &\geq \alpha\|u\|^2 \, . \end{aligned}$$

Sei X_h ein Teilraum von X.
Die Gleichungen

$$a(u,v) = < b, v > \, , \quad \forall \, v \in X$$

und

$$a(u_h, v_h) = < b, v_h > \, , \quad \forall \, v_h \in X_h$$

haben eindeutige Lösungen u bzw. u_n (Lax-Milgram).

Es gilt:

$$\|u - u_n\| \leq \frac{M}{\alpha} \inf_{v_n \in X_h} \|u - v_n\| \ .$$

<u>Beweis:</u> Für jedes $v_h \in X_h$ gilt:

$$\begin{aligned}
\alpha \|u - u_n\|^2 &\leq a(u - u_n, u - u_n) \\
&= a(u - u_n, u - v_h) + a(u - u_n, v_n - u_n) \\
&= a(u - u_n, u - v_h) + a(u, v_h - u_n) - a(u_n, v_h - u_h) \\
&= a(u - u_n, u - v_h) + < b, v_h - u_n > - < b, v_h - u_n > \\
&= a(u - u_n, u - v_h) \\
&\leq M \|u - u_n\| \cdot \|u - v_h\| \ .
\end{aligned}$$

Literatur

Adams, R.A.: Sobolev Spaces. New York: Academic Press, 1975.

Ciarlet, P.G.: The Finite Element Method for Elliptic Problems. North-Holland, 11978.

Gilbarg, D., **Trudinger**, N.S.: Elliptic Partial Differential Equations of Second Order. Springer 1977.

Hackbusch, W.: Theorie und Numerik elliptischer Differentialgleichungen. Teubner 1986.

Heuser, H.: Funktionalanalysis. Teubner 1986.

Kelley, J.L. **Namioka**, I.: Linear Topological Spaces. Springer 1963.

Morrey, C.B. Jr.: Multiple Integrals in the Calculus of Variations. Springer 1966.

Natanson, I.P.: Theorie der Funktionen einer reellen Veränderlichen.

Robertson, A.P., **Robertson**, W.J.: Topological Vector Spaces. Cambridge 1964.

Smirnow, W.I.: Lehrgang der Höheren Mathematik, Teil V, VEB Deutscher Verlag der Wissenschaften 1962.

Sobolev, S.L.: Applications of Functional Analysis in Mathematical Physics. American Mathematical Society 1963.

Schwarz, H.R.: Methode der finiten Elemente. Teubner 1980.

Strang, G.J.: Fix: An Analysis of the Finite Element Method. Prentice-Hall 1973.

Yosida, K.: Functional Analysis, 2nd edition. Springer 1968.

Zienkiewicz, O.C.: The Finite Element Method in Engineering Science. Second edition, McGraw-Hill 1971. (Es gibt eine neuere Ausgabe.)