TECHNISCHE MECHANIK 7(1986) Heft 2
Manuskripteingang: 15. 7. 1985

Moderne Methoden zur Behandlung von Konvektions-Diffusions-Gleichungen

Hans-Görg Roos

1. Einleitung

Wir betrachten das Problem der numerischen Lösung der Konvektions-Diffusions-Gleichung

$$-\epsilon\Delta u + b(x)\cdot\nabla u + c(x)u = f(x) \text{ in } \Omega\subset R^2 \qquad (1)$$
$$u = 0 \quad \text{auf } \partial\Omega .$$

Es gibt eine ganze Reihe von Aufgaben, die zu Randwertproblemen vom Typ (1) führen, z. B. die Berechnung von Stoff- und Wärmetransport von Grundwasserströmungen in geologischem Untergrund, die Berechnung von Temperaturfeldern beim MAG- bzw. Aluminiumschweißen, die Berechnung des Temperaturfeldes im Ziehstein beim Drahtziehen.

Betrachtet man bei dem zweiten genannten Problemkreis konkret die Berechnung des stationären Temperaturfeldes einer ebenen bewegten Platte mit einer Punktquelle, so genügt die Temperatur T der Fourier-Kirchhoffschen Differentialgleichung

$$-\frac{1}{\rho c_p}\left[\lambda_{xx}\frac{\partial^2 T}{\partial x^2} + \lambda_{yy}\frac{\partial^2 T}{\partial y^2}\right] + v_x\frac{\partial T}{\partial x} + v_y\frac{\partial T}{\partial y} = \frac{\tilde{\dot{e}}_i}{\rho c_p},$$

dabei sind ρc_p die Wärmekapazitätsdichte, λ_{xx}, λ_{yy} Wärmeleitkoeffizienten, v_x, v_y die Geschwindigkeitskomponenten des Transportes, $\dot{e}_i$ die innere Energieumsetzung. Die Randbedingungen an der Modelleintrittsgrenze berücksichtigen die Zuführung des Enthalpiestromes gemäß

$$v_x\rho c_p(T-t_o) = \alpha_{\ddot{a}qu.}(T-t_o),$$

dabei ist t_o die Temperatur der Platte an der Modellgrenze, $\alpha_{\ddot{a}qu.}$ ein Wärmeübertragungskoeffizient, die übrigen Randbedingungen sind natürliche (weitere Einzelheiten findet man in [5]). Definiert man eine Peclet-Zahl Pe (das entspricht ϵ^{-1}) gemäß $Pe = \frac{1}{\lambda}\rho c_p v_x L$ mit einer für die Platte charakteristischen Länge L, so liegt Pe für die in [5] behandelten Probleme in etwa im Bereich $50 \leqslant Pe \leqslant 100$.

Es sei h eine charakteristische Länge des Diskretisierungsverfahrens zur Lösung von (1). Dann ist bekannt, daß, sobald eine lokale Peclet-Zahl h/ϵ einen gewissen kritischen Wert überschreitet, es sowohl nicht möglich ist, die konvektiven Terme auf der Basis eines vorhandenen Poisson-Lösers iterativ zu berücksichtigen, als auch Standard-Diskretisierungsverfahren direkt auf (1) anzuwenden. Bei der Anwendung von Standard-Diskretisierungsverfahren kommt es oft zu unerwünschten Oszillationen, und man kann auch Beispiele dafür angeben, daß eine lokale Netzverfeinerung in Bereichen von Oszillationen nicht notwendig sinnvoll ist, sondern a-priori Informationen über die Gebiete voraussetzt, die diese Oszillationen hervorrufen.

Mit dem Programmsystem T 85 [5] wurde bewiesen, daß ein Finites-Element-Verfahren (FEM) vom upwind-Typ für die Berechnung stationärer zweidimensionaler Temperaturfelder für konvektiv dominante Vorgänge unter Berücksichtigung thermischer Randbedingungen, Energieumsetzungen und zu- bzw. abgeführter Enthalpieströme insbesondere beim MAG-Schweißen geeignet ist. Nun wurden in den letzten 10 Jahren zahlreiche Vorschläge für upwind-FEM unterbreitet (vgl. die Literaturangaben in [7]); Anliegen dieses Artikels ist es, zwei der im Mittelpunkt der Diskussion stehenden Verfahren vorzustellen, und zwar eine hybride upwind-FEM-Technik und die Stromliniendiffusionsmethode. Im Fall extrem großer lokaler Peclet-Zahlen verweisen wir auf die exponentiell-angepaßten FEM [9].

Wir beschränken uns hier auf den stationären Fall; es ist durchaus möglich, die vorzustellenden Methoden auf instationäre Fälle einschließlich komplizierterer Randbedingungen auszudehnen [1], [4], [10].

2. Fehlerabschätzungen und inverse Isotonie

Ω sei ein zweidimensionales Gebiet, konvex und polygonal. Die Daten von (1) seien glatt, zudem sei $c \geqslant 0$, $c - \frac{1}{2}\operatorname{div} b \geqslant 0$.

Es sei $L_2(\Omega)$ die Menge aller Funktionen über Ω, die quadratisch integrierbar sind, ausgestattet mit der Norm

$$\|v\|_o = \left(\int_\Omega v^2\, d\Omega\right)^{\frac{1}{2}}.$$

Die Menge aller Funktionen aus dem $L_2(\Omega)$, deren Ableitungen ebenfalls zum $L^2(\Omega)$ gehören und die zudem auf dem Rand von Ω verschwinden, bezeichnen wir mit $\overset{\circ}{H}{}^1(\Omega)$ und verwenden die Norm

$$\|v\|_1 = \left(\int_\Omega |\nabla v|^2\, d\Omega\right)^{\frac{1}{2}}.$$

Ω werde wie bei der Methode der finiten Elemente üblich, in Dreiecke zerlegt (wir beschränken uns auf diesen Fall), die Zerlegung sei regulär in dem Sinne, daß der kleinste vorkommende Winkel gleichmäßig nach unten beschränkt ist. Der Parameter h sei der maximale Durchmesser aller Dreiecke der Zerlegung. Die Ansatzfunktionen für eine FEM zur Lösung von (1) wählen wir aus der Menge $V_h \subset \overset{\circ}{H}{}^1(\Omega)$ wie üblich. Führt man dann noch die Abkürzung

$$\int_\Omega v\, w\, d\Omega = (v, w)$$

ein, so kann man die normale FEM beschreiben durch: Gesucht ist ein $u_h \in V_h$ mit der Eigenschaft

$$\epsilon(\nabla u_h, \nabla w_h) + (b\nabla u_h + c u_h, w_h)$$

$$= (f, w_h) \text{ für alle } w_h \in V_h . \qquad (2)$$

Fehlerabschätzungen für die FEM sind wohlbekannt, als Maß für den Fehler verwendet man oft die L^2- oder $\overset{\circ}{H}{}^1$-Norm von $u-u_h$, gebräuchlich ist aber auch das wesentliche Maximum von $u-u_h$, das wir hier mit $\|u-u_h\|$ bezeichnen. Mit C bezeichnen wir (verschiedene) Konstanten, die von h und ϵ unabhängig sind.

Im Fall nichtdominanter Konvektion (den wir durch $\epsilon = 1$ kennzeichnen) gilt für lineare Dreieckelemente

$$\|u-u_h\|_1 \leqslant Ch, \quad \|u-u_h\|_0 \leqslant Ch^2, \quad \|u-u_h\| \leqslant Ch^2\, |\ln h| .$$

Im Fall dominanter Konvektion dagegen kann man Konvergenz der FEM bei linearen Dreieckelementen nur für kleine lokale Peclet-Zahlen mit $h/\epsilon \leqslant 1$ beweisen, wobei

$$\|u-u_h\|_1 \leqslant Ch, \quad \|u-u_h\|_0 \leqslant Ch^2/\sqrt{\epsilon} .$$

Da man die Bedingung $h/\epsilon \leqslant 1$ oft nicht realisieren kann, ist die normale FEM im Fall dominanter Konvektion nicht geeignet.

Es gibt einen weiteren Grund, der bei konvektionsdominanten Problemen gegen die normale FEM spricht. Die Lösung von (1) genügt nämlich einem Maximumprinzip. So folgt aus

$$(\epsilon\nabla w_1, \nabla w) + (b\nabla w_1 + c w_1, w) \leqslant (\epsilon\nabla w_2, \nabla w)$$
$$+ (b\nabla w_2 + c w_2, w)$$

für alle nichtnegativen w des $\overset{\circ}{H}{}^1(\Omega)$ die Relation

$$w_1 \leqslant w_2,$$

und dies sichert z. B. eine gewisse Vorzeicheninvarianz von u. Bei der Diskretisierung mit einer normalen FEM geht im konvektionsdominanten Fall diese Eigenschaft aber im allgemeinen verloren!

Matrizen A, die solch eine Vorzeicheninvarianz sichern, heißen invers-isoton (aus $Ax_1 \leqslant Ax_2$ folgt für die Vektoren x_1, x_2 die Relation $x_1 \leqslant x_2$). Eine spezielle Klasse von invers-isotonen Matrizen sind die M-Matrizen, dies sind invers-isotone Matrizen mit nichtpositiven Außerdiagonalelementen. Im Fall nichtpositiver Außerdiagonalelemente ist nämlich eine Matrix genau dann inversisoton, wenn ein positiver Vektor z (ein majorisierendes Element) existiert mit $Az > 0$. Leider ist in anderen Fällen eine so schöne Charakterisierung der inversen Isotonie nicht möglich [3].

Hat man nur ein schwach majorisierendes Element ($z \geqslant 0$ mit $Az \geqslant 0$), so folgt die inverse Isotonie bei nichtpositiven Außerdiagonalelementen, wenn die Ketteneigenschaft vorliegt. Diese ist folgendermaßen definiert: N_1 sei die Indexmenge, gekennzeichnet durch $(Az)_k = 0$, N_2 die Indexmenge, gekennzeichnet durch $(Az)_k > 0$ ($(Az)_k$ sei die k-te Komponente des Vektors Az). Dann besitzt A die Ketteneigenschaft, wenn für alle $i \in N_1$ endlich viele Indizes $i_o, i_1, \ldots, i_r$ existieren, so daß die Elemente a_{i_{k-1}, i_k} der Matrix A von Null verschieden sind und $i_o = i$, $i_r \in N_2$.

Untersucht man ein Gleichungssystem

$$Ay = g$$

mit invers-isotonem A, so ist einmal die eindeutige Lösbarkeit gesichert, zudem folgt aus $g \leqslant 0$ die Relation $y \leqslant 0$, und es existiert eine Konstante C mit

$$|y| \leqslant C\, |g| \qquad (3)$$

($|y|$ sei das Maximum der Beträge der Komponenten des Vektors y). Diese Beziehung ermöglicht Stabilitätsuntersuchungen.

3. Hybride upwind-FEM

Wir setzen grundsätzlich in diesem Abschnitt voraus, daß die Vernetzung vom schwach spitzen Typ sei, unter den Innenwinkeln aller Dreiecke also keine stumpfen Winkel auftreten. Bekanntlich diskretisiert dann das klassische FEM-Verfahren mit linearen Dreieckelementen den Operator $-\Delta$ so, daß für die das diskrete Problem erzeugende Matrix A_h gilt $(A_h)_{i,j} \leqslant 0$ für $i \neq j$ und A_h sogar M-Matrix ist.

Die Grundidee der hybriden upwind-FEM besteht nun darin, die restlichen Terme in Anlehnung an die Integralbilanztechnik so zu diskretisieren, daß diese Eigenschaft der Matrix A_h für alle Parameterkonstellationen zwischen ϵ und h nicht zerstört wird.

Zu jedem Knoten P_i der Dreieckszerlegung wird zunächst ein duales Gebiet D_i konstruiert. D_i wird z. B. begrenzt von den Abschnitten der Mittelsenkrechten der zu P_i benachbarten Dreiecke (man könnte die Konstruktion auch mit den Seitenhalbierenden durchführen). Es sei $\Lambda_i = \{ j \mid$ es existiert ein Dreieck, zu dem P_i und P_j gehören $\}$, Λ_i ist also die Indexmenge der „benachbarten" Knoten. P_{ij} sei der Mittelpunkt von $\overline{P_i P_j}$ und Γ_{ij} die Vereinigung der P_{ij} enthaltenden geradlinigen Abschnitte vom Rand ∂D_i von D_i.

Zur speziellen Diskretisierung des Konvektionstermes gehen wir aus von der Aufspaltung

$$(b\nabla u_h, w_h) = (\operatorname{div}(u_h b), w_h) - ((\operatorname{div} b)\, u_h, w_h) .$$

Der erste Summand wird folgendermaßen behandelt:

$$(\operatorname{div}(u_h b), w_h) = \sum_i \int_{D_i} \operatorname{div}(u_h b)\, w_h\, dx$$

$$\approx \sum_i w_h(P_i) \int_{D_i} \operatorname{div}(u_h b)\, dx ,$$

der Gaußsche Integralsatz liefert

$$(\operatorname{div}(u_h b)\, w_h) \approx \sum_i w_h(P_i) \int_{\partial D_i} b \cdot \nu(\partial D_i)\, u_h\, d\partial D_i$$

$$= \sum_i w_h(P_i) \sum_{j\in\Lambda_i} \int_{\Gamma_{ij}} (b \cdot \nu_{ij})\, u_h\, d\Gamma_{ij} .$$

Die Anwendung der Integrationsformel

$$\int_{\Gamma_{ij}} (b \cdot \nu_{ij})\, u_h\, d\Gamma_{ij}$$

$$\approx (b(P_{ij}) \cdot \nu_{ij}\ \operatorname{meas}(\Gamma_{ij})(\lambda_{ij}\, u_h(P_i)$$

$$+ (1-\lambda_{ij})\, u_h(P_j))$$

liefert

$$(\operatorname{div}(u_h b), w_h)$$
$$\approx \sum_i w_h(P_i) \sum_{j\in\Lambda_i} (b(P_{ij})\cdot \nu_{ij})\,\mathrm{meas}(\Gamma_{ij})(\lambda_{ij}\, u_h(P_i) + (1-\lambda_{ij})\, u_h(P_j))\,.$$

Ähnlich wird der zweite Summand der obigen Aufspaltung approximiert:

$$((\operatorname{div} b)\, u_h, w_h) = \sum_i \int_{D_i} (\operatorname{div} b)\, u_h\, w_h\, dx$$
$$\approx \sum_i u_h(P_i)\, w_h(P_i) \int_{D_i} \operatorname{div} b\, dx\,,$$

also

$$((\operatorname{div} b)\, u_h, w_h)$$
$$\approx \sum_i u_h(P_i)\, w_h(D_i) \sum_{j\in\Lambda_i} (b(P_{ij})\cdot \nu_{ij})\,\mathrm{meas}(\Gamma_{ij})\,.$$

Insgesamt lautet die Approximation des Konvektionstermes

$$(b\nabla u_h, w_h)$$
$$\approx \sum_i w_h(P_i) \sum_{j\in\Lambda_i} (b(P_{ij})\cdot \nu_{ij})\,\mathrm{meas}(\Gamma_{ij})\,((\lambda_{ij}-1)\, u_h(P_i) + (1-\lambda_{ij})\, u_h(P_j))\,.$$

Wenn wir mit Φ_k die dem Knoten P_k zugeordnete Basisfunktion bezeichnen, so erzeugt diese Approximation eine Matrix B_h, deren Element in der k-ten Zeile und l-ten Spalte wir berechnen können, indem wir setzen $u_h := \Phi_l$, $w_h := \Phi_k$. Dann gilt

$$(B_h)_{kk} = \sum_{j\in\Lambda_k} (b(P_{kj})\cdot \nu_{kj})\,\mathrm{meas}\,\Gamma_{kj}\,(\lambda_{kj}-1)$$
$$(B_h)_{k,l} = (b(P_{kl})\cdot \nu_{kl})\,\mathrm{meas}(\Gamma_{kl})(1-\lambda_{kl}), \text{ wenn } l\in\Lambda_k$$
$$(B_h)_{k,l} = 0 \qquad\qquad \text{sonst.}$$

Mit der Wahl

$$\lambda_{kl} = \begin{cases} 1, & \text{wenn } b(P_{kl})\cdot \nu_{kl} \geqslant 0 \\ 0, & \text{wenn } b(P_{kl})\cdot \nu_{kl} < 0 \end{cases}$$

wird gesichert, daß die Matrix B_h nichtpositive Außerdiagonalelemente besitzt.

Zur Diskretisierung des Absolutgliedes von (1) nützen wir eine „mass lumping"-Technik, die Diskretisierung erzeugt eine Diagonalmatrix C_h mit

$$(C_h)_{k,l} = \begin{cases} c(P_k)\,\mathrm{meas}\, D_k & \text{für } l = k \\ 0 & \text{für } l \neq k \end{cases}.$$

Faßt man die Werte von u_h in den Knoten P_k zu einem Vektor v_h zusammen, so erzeugt die hybride upwind-Technik das Gleichungssystem

$$L_h v_h \equiv (\epsilon A_h + B_h + C_h) v_h = f_h\,. \tag{4}$$

Wesentlich ist nun, daß die Matrix L_h eine M-Matrix ist. Dies folgt daraus, daß die Matrix L_h nichtpositive Außerdiagonalelemente besitzt, die Zeilensummen nichtnegativ sind und L_h die Ketteneigenschaft besitzt.

Als Beispiel betrachten wir die Gleichung

$$-\epsilon\Delta u + p\,\frac{\partial u}{\partial x_1} + q\,\frac{\partial u}{\partial x_2} = f(x_1, x_2) \tag{5}$$

im Einheitsquadrat mit positiven Konstanten p und q. Wir legen eine Vernetzung vom Friedrichs-Keller-Typ zugrunde, in einem regelmäßigen Quadratnetz werde jedes Quadrat in zwei Dreiecke zerlegt durch eine Gerade der Form $x_2 = x_1 + \delta$.

Sei v_{nm} der Funktionswert von u_h im Knoten mit den Koordinaten $x_1 = n\cdot h$, $x_2 = mh$, ferner $E = [e_{\nu\mu}]_{\nu,\mu = -1,0,1}$ eine Matrix vom Format 3 x 3. Wir schreiben dann (4) in unserem Spezialfall als Differenzengleichung

$$\sum_{\nu,\mu=-1,0,1} e_{\nu\mu}\, v_{n+\nu,m+\mu} = (f_h)_{n+m}$$

und kennzeichnen diesen Neun-Punkte-Differenzenstern im folgenden durch die Angabe der Matrix E.

Verwendet man zur Konstruktion der dualen Bereiche Mittelsenkrechten, so erhält man

$$E = \epsilon \begin{pmatrix} \cdot & -1 & \cdot \\ -1 & 4 & -1 \\ \cdot & -1 & \cdot \end{pmatrix} + h \begin{pmatrix} \cdot & \cdot & \cdot \\ -p & p+q & \cdot \\ \cdot & -q & \cdot \end{pmatrix}$$

(ein $\cdot$ bedeutet eine Null). Das Ergebnis ist also identisch mit dem üblichen upwind-Differenzenverfahren. Die hybride upwind-Technik besitzt den Vorteil ihrer Anwendbarkeit auch bei komplizierterer Geometrie von Ω.

Als nächstes stellt sich die Frage, ob die Lösung des diskreten Problems gleichmäßig stabil ist in dem Sinne, daß man für die Lösung von (4) eine Abschätzung vom Typ (3) mit von ϵ unabhängigem C angeben kann. Dies wäre dann der Fall, wenn man ein majorisierendes Element z fände mit $z \geqslant c_0 > 0$, $Az \geqslant c_1 > 0$ mit von ϵ unabhängigen Konstanten c_0, c_1. Ist der Koeffizient des Absolutgliedes in (1) $c(x) > 0$, so ist die gleichmäßige Stabilität gesichert. In anderen Fällen gelingt es oft, glatte Funktionen $z(x)$ zu finden, so daß $z(x) \geqslant c_0 > 0$ und

$$-\epsilon\Delta z + b\cdot\nabla z + cz \geqslant c_1 > 0.$$

Die Schwierigkeiten in der Konstruktion eines majorisierenden Elementes für das diskrete Problem bei bekanntem majorisierenden Element $z(x)$ für das stetige Problem liegen darin begründet, daß L_h im allgemeinen keine konsistente Approximation des ursprünglichen elliptischen Operators im Sinne der Theorie von Differenzenverfahren ist. Das liegt aber nicht an der hybriden upwind-Technik, selbst $-\Delta u$ wird mit der üblichen FEM-Technik nur für spezielle Dreiecksnetze oder spezielle Funktionen u konsistent approximiert. Seien etwa $(x_i + p_j, y_i + q_j)$ mit $j = 1, \ldots, n$ die kartesischen Koordinaten von $P_j \in \Lambda_i$ und

$$\delta_j = -\frac{1}{2}\left(\frac{p_{j+1}(p_{j+1}-p_j) + q_{j+1}(q_{j+1}-q_j)}{p_j q_{j+1} - q_j p_{j+1}} + \frac{p_{j-1}(p_{j-1}-p_j) + q_{j-1}(q_{j-1}-q_j)}{p_{j-1} q_j - p_j q_{j-1}}\right).$$

Dann approximiert A_h den Operator $-\Delta$ genau dann für beliebige glatte Funktionen u konsistent, wenn

$$\sum_{j=1}^{n} \delta_j\, p_j q_j = 0, \quad \sum_{j=1}^{n} \delta_j (p_j^2 - q_j^2) = 0 \quad \text{und}$$

$$\sum_{j=1}^{n} \delta_j p_j^2 \neq 0 .$$

Auf weniger regelmäßigen Netzen sind diese Beziehungen im allgemeinen nicht erfüllt.

Versteht man unter einer regelmäßigen Vernetzung von Ω eine Zerlegung, bei der sich je zwei Dreiecke durch Drehung um 180° oder Verschiebung ineinander überführen lassen, so ist im Fall einer regelmäßigen Vernetzung die Konsistenz auch der hybriden upwind-Technik gesichert, und die Existenz eines majorisierenden Elementes $z(x)$ für die Differentialgleichung (1) sichert dann die gleichmäßige Stabilität. Man kann damit erwarten, daß im größten Teil des Gebietes (Grenzschichtbereiche sind natürlich ausgeschlossen) die Lösung stabil von erster Ordnung approximiert wird.

Tatsächlich kann man Konvergenzresultate in der $\|\cdot\|$-Norm unter Ausnutzung der inversen Isotonie des diskreten Problems beweisen. Während in [1] die Abhängigkeit der vorkommenden Konstanten vom Parameter ϵ nicht diskutiert wird, findet man präzisere Aussagen in dieser Richtung in [6].

Im eindimensionalen Fall entspricht bei der angegebenen Parameterwahl für die λ_{kl} das Verfahren dem üblichen upwind-Differenzenverfahren. Die getroffenen Aussagen über die stabile Approximierbarkeit der Lösung von der Ordnung Eins sind dann wohlbekannt. Im zweidimensionalen Fall werden in [1] an fünf Beispielen verschiedene Varianten dieser upwind-Technik getestet. Ein Beispiel ist die Randwertaufgabe

$$-\epsilon \Delta u + \frac{\partial u}{\partial x} + \frac{\partial u}{\partial y} = 0 \quad \text{in} \quad \Omega$$

mit $\Omega = (x,y) \mid 0 < x < I$ und $0 < y < \frac{\sqrt{3}}{2}$ und solchen Randbedingungen, daß

$$u(x,y) = \frac{\exp(\frac{x}{\epsilon} - 1)}{\exp(\frac{1}{\epsilon} - 1)} \left\{ 1 + \frac{\exp(\frac{y}{\epsilon}) - 1}{\exp(\frac{\sqrt{3}}{2\epsilon}) - 1} \right\}$$

die Randwertaufgabe löst; für ϵ wird $\epsilon = \frac{1}{8}, \frac{1}{16}, \frac{1}{32}, \frac{1}{64}$ gewählt.

Das Gebiet wird in $(m_D)^2$ kongruente Rechtecke zerlegt, diese jeweils in zwei Dreiecke, es sei $m_D = 4, 8, 16$. Es zeigt sich, daß der Fehler $\|u - u_h\|$ unabhängig von den oben angegebenen ϵ-Werten in der Größenordnung von 10^{-I} liegt, was bei der hier vorliegenden groben Vernetzung durchaus die theoretischen Aussagen bestätigt.

4. Die Methode der Stromliniendiffusion

Schon seit einigen Jahren bekannt ist der Vorschlag, eine normale FEM zu stabilisieren durch die Verwendung von Testfunktionen, die von den Ansatzfunktionen verschieden sind (Petrov-Galerkin-FEM). Zunächst wurden als Testfunktionen polynomiale Störungen der Ansatzfunktionen verwendet. Die Methode der Stromliniendiffusion basiert auf Testfunktionen vom Typ

$$w_h + \beta b \cdot \nabla w_h \quad \text{mit} \quad w_h \in V_h .$$

β ist ein später geeignet zu wählender Parameter. Wählt man in der Gleichung (2) eine Testfunktion von diesem Typ, so ist aber für die üblicherweise benutzten Ansatzfunktionen, die nur stetig sind beim Übergang von einem Element zum anderen, das in der Gleichung (2) auftretende Integral

$$\int_\Omega \nabla u_h \, \nabla (b \cdot \nabla w_h) \, d\Omega$$

nicht erklärt. Wie immer bei nichtkonformen Methoden hilft man sich, indem man zunächst nicht erklärte Integrale auf geeignete Weise durch die Summe von Integralen über die Elemente ersetzt. Die Methode der Stromliniendiffusion ist letztlich gekennzeichnet durch:
Gesucht ist ein $u_h \in V_h$ mit

$$\begin{aligned} &\epsilon(\nabla u_h, \nabla w_h) - \epsilon \sum_T (\nabla u_h, \beta b \nabla w_h)_T \\ &+ (b \nabla u_h + c u_h, w_h + \beta b \nabla w_h) \\ &= (f, w_h + \beta \nabla w_h) \quad \text{für alle } w_h \in V_h . \end{aligned} \tag{6}$$

Dabei ist die Summation über alle Dreiecke der Zerlegung zu erstrecken und

$$(\Delta u_h, b \nabla w_h)_T = \int_T (\Delta u_h)(b \nabla w_h) \, dT .$$

Die folgende Tabelle kennzeichnet die Größenordnung der auftretenden Terme:

Terme:	Summand	Größenordnung
(1)	$\epsilon(\nabla u_h, \nabla w_h)$	$0(\epsilon)$
(2)	$\epsilon \sum_T (\Delta u_h, \beta b \nabla w_h)_T$	$0(\frac{\beta\epsilon}{h})$
(3)	$(b \nabla u_h, w_h)$	$0(h)$
(4)	$(b \nabla u_h, \beta b \nabla w_h)$	$0(\beta)$
(5)	$(c u_h, w_h)$	$0(h^2)$
(6)	$(c u_h, \beta b \nabla w_h)$	$0(\beta h)$
(7)	(f, w_h)	$0(h^2)$
(8)	$(f, \beta \nabla w_h)$	$0(\beta h)$

Allein diese Angaben lassen es sinnvoll erscheinen, β proportional zu h zu wählen. Eine theoretische Fehleranalyse [4] führt in der Tat zu

$$\beta = \begin{cases} \beta^* h, & \text{wenn } \epsilon < h \\ 0, & \text{wenn } \epsilon \geqslant h \end{cases} . \tag{7}$$

Wir nehmen nun an, wir würden im Fall $\epsilon < h$ bei der Wahl von β gemäß (7) lineare (oder bilineare im Fall einer Rechtecksvernetzung) Ansatzfunktion wählen. Dann ist der zweite Summand identisch Null, und die wesentliche Abweichung der Methode der Stromliniendiffusion gegenüber der üblichen FEM besteht darin, daß folgende Ersetzung des Konvektionstermes realisiert wird:

$$(b \nabla u_h, w_h) := (b \nabla u_h, w_h) + \beta^* h (b \nabla u_h, b \nabla w_h). \tag{8}$$

Schreibt man in dieser Beziehung

$$\beta^* h (b \nabla u_h, b \nabla w_h) = \beta^* h (\nabla u_h, b^T b \nabla w_h) \tag{9}$$

(b^Tb ist die Matrix des dynamischen Produkts b^Tb), so erkennt man die enge Beziehung der Methode der Stromliniendiffusion mit der Methode der künstlichen Diffusion in Stromrichtung [2].
Bei dieser Methode wird ein lokales Koordinatensystem (s, n) parallel und senkrecht zum Geschwindigkeitsvektor b eingeführt, nach der Transformation die künstliche Diffusion gemäß

$$k'_{sn} = \begin{bmatrix} k' & 0 \\ 0 & 0 \end{bmatrix}, \quad k' = \frac{1}{2}\alpha^* |b| h$$

eingeführt und dann zurücktransformiert und die übliche FEM-Technik angewandt. Dies entspricht einer Ersetzung des Diffusionstermes gemäß

$$(\nabla u_h, \nabla w_h) := (\nabla u_h, \nabla w_h) + \frac{k'}{|b|^2}(\nabla u_h, b^T b \nabla w_h) \tag{10}$$

Ein Vergleich von (8), (9) und (10) zeigt, daß bei entsprechender Abstimmung von α^* und β^* die Zusatzterme übereinstimmen.
Wir betrachten als Beispiel wieder die Gleichung (5) bei der gleichen Zerlegung von Ω wie oben. Die Methode der Stromliniendiffusion mit linearen Dreieckelementen liefert

$$E = \epsilon \begin{pmatrix} \cdot & -1 & \cdot \\ -1 & 4 & -1 \\ \cdot & -1 & \cdot \end{pmatrix} + \frac{h}{6}\begin{pmatrix} \cdot & -p+2p & p+q \\ -2p+q & \cdot & 2p-q \\ -(p+q) & p-2q & \cdot \end{pmatrix}$$

$$+ \beta^* h \begin{pmatrix} \cdot & pq-q^2 & -pq \\ pq-p^2 & 2(p^2+q^2-pq & pq-p^2) \\ -pq & pq-q^2 & \cdot \end{pmatrix} .$$

Im Fall bilinearer Ansatzfunktionen ergibt sich

$$E = \frac{\epsilon}{3}\begin{pmatrix} -1 & -1 & -1 \\ -1 & 8 & -1 \\ -1 & -1 & -1 \end{pmatrix} + \frac{h}{12}\begin{pmatrix} -p+q & 4q & p+q \\ -4p & \cdot & 4p \\ -p-q & -4q & p-q \end{pmatrix}$$

$$+\frac{\beta^* h}{6}\begin{pmatrix} -p^2-q^2+3pq & 2p^2-4q^1 & -p^2-q^2-3pq \\ -4p^2+2q^2 & 8p^2+8q^2 & -4p^2+2q^2 \\ -p^2-q^2-3pq & 2p^2-4q^2 & -p^2-q^2+3pq \end{pmatrix}$$

Leider kann man den Parameter β^* nicht so wählen, daß M-Matrizen entstehen.

Bei gewissen Glattheitsvoraussetzungen an die Lösung kann man [4] bei der Verwendung von Polynomen k-ten Grades als Ansatzfunktionen bei einer Dreieckszerlegung den Fehler abschätzen zu ($\epsilon < h$)

$$\|u - u_h\|_0 \leqslant C^* h^{k+1/2}$$
$$\|u - u_h\|_1 \leqslant C^* h^{k+1/2}/\epsilon$$
$$\|b \cdot \nabla (u - u_h)\|_0 \leqslant C^* h^k .$$

Die Abschätzung für den Fehler der Ableitung in Stromrichtung besitzt die optimal erreichbare Ordnung. Bei einer regelmäßigen Vernetzung vom Friedrichs-Keller-Typ und linearen Dreieckelementen (k = 1) läßt sich die erste Abschätzung verbessern zu

$$\|u - u_h\|_0 \leqslant C^* h^2.$$

Die Konstanten C^* in diesen Abschätzungen sind aber nicht von ϵ unabhängig, sondern besitzen mit von ϵ und h unabhängigen Konstanten C die Form

$C^* = C |u|_{k+1}$, dabei ist

$$|u|_{k+1} = \left\{ \int_\Omega \sum_{l_1+l_2=k+1} \left(\frac{\partial^{k+1} u}{\partial^{l_1} x_1 \, \partial^{l_2} x_2}\right)^2 d\Omega \right\}^{\frac{1}{2}}$$

durchaus von ϵ abhängig.

Wesentlich ist aber, daß es möglich ist, den Fehler in bestimmten Teilgebieten so abzuschätzen, daß nur Integrale über gewisse andere Teilgebiete auftreten. Dies zeigt, daß in Nichtgrenzschichtbereichen C^* unabhängig von ϵ ist und somit in diesen Bereichen die Methode der Stromliniendiffusion auch für extrem große lokale Pecletzahlen gute Ergebnisse liefert. In den Grenzschichtbereichen muß man in solchen Fällen auf exponentiell angepaßte Methoden zurückgreifen.
In [4] wird an verschiedenen Testbeispielen die numerische Konvergenzrate des Verfahrens untersucht. Im stationären Fall wird dazu betrachtet

$$-\epsilon \Delta u + (1.1 - y)\frac{\partial u}{\partial x} + (x - 0.1)\frac{\partial u}{\partial y} + u = f$$

mit $\epsilon = 10^{-6}$ · f wird so gewählt, daß für $u_o = x^3 + y^3$ gilt

$$(1.1 - y)\frac{\partial u_o}{\partial x} + (x - 0.1)\frac{\partial u_o}{\partial y} + u_o = f .$$

Das Gebiet Ω wird begrenzt von $x = 0.2$ mit $0 \leqslant y \leqslant 1$, $y = 0$ mit $0.2 \leqslant x \leqslant 1$, $y = 1$ mit $0.2 \leqslant x \leqslant 0.661\,438$, $x = 1$ mit $0 \leqslant y \leqslant 0.25$ und dem Kreisbogen um (0.0, 0.25), der (1.0, 0.25) und (0.661 438,1) verbindet.
Die Randbedingungen werden gewählt zu

$$u = \begin{cases} x^3 + y^3 & \text{für } x = 0.2 \text{ mit } 0 \leqslant y \leqslant 1 \text{ und } y = 0 \text{ mit } 0.2 \leqslant x \leqslant 1 \\ 0 & \text{auf dem übrigen Teil des Randes} \end{cases}$$

Bei dieser Wahl verhält sich u für $\epsilon \to 0$ im wesentlichen wie u_o, deshalb wird die Güte der Approximation an der Differenz $u_o - u_h$ gemessen.
Das Gebiet wird mit einer Ausgangsschrittweite von $h = \frac{1}{8}$ in Dreiecke zerlegt, am krummlinigen Teil des Randes wird wie üblich modifiziert. Es wird $\beta = \beta^* h$ gewählt mit

$$\beta^* = \frac{1}{2}\frac{|b_1| + |b_2|}{|b|^2}$$

Das normale FEM-Verfahren ($\beta^* = 0$) zeigt für $h = \frac{1}{8}, \frac{1}{16}, \frac{1}{24}, \frac{1}{32}$ keinerlei Anzeichen der Konvergenz, dagegen erhält man für das Stromliniendiffusionsverfahren die Konvergenzraten 0.48, 0.49, 0.51 in der Norm $L^2(\Omega)$, 2.4, 2.2, 2.0 in der Norm des $L^2(\Omega)$ und 1.1, 1.0, 1.0 in der Norm $\|b \cdot \nabla(u_o - u_h)\|_{L^2(\Omega')}$ $(\Omega' \subset\subset \Omega)$.
Dies entspricht den theoretischen Konvergenzraten.

LITERATUR

[1] Ikeda, T.: Maximum principle in finite element models for convection-diffusion phenomena. Amsterdam 1983.

[2] Kelly, D. W., Nakazawa, S., Zienkiewicz, O. C.: A note on upwinding and anisotropic balancing dissipation in finite element approximations to convective diffusion problems. Int. J. Num. Meth. in Engg. 15, 1705 – 1711 (1980).

[3] Lorenz, J.: Zur Inversmonotonie diskreter Probleme. Num. Math. 27, 227 – 238 (1977).

[4] Nävert, U.: A finite element method for convection-diffusion problems. Diss., Göteborg 1982.

[5] Penzel, J.: Forschungsbericht (unveröffentlichtes Manuskript), TU Dresden 1984.

[6] Risch, U.: Hybride upwind-FEM für schwach gekoppelte elliptische Systeme (in Vorbereitung).

[7] Roos, H.-G.: Necessary convergence conditions for upwind schemes in the two-dimensional case. Preprint 18/1983, IMATH Berlin.

[8] Roos, H.-G., Risch, U.: Hybride upwind-Schemata. Preprint 07–04–85, TU Dresden 1985.

[9] Schieweck, F.: Finite element methods for singularly perturbed partial differential equations. In: „Singularly perturbed differential equations and applications" (ed. J. Foerste). IMECH Berlin 1985.

[10] Windisch, G.: Konservative Differenzenapproximation von Diffusions-Konvektionsproblemen auf instationären Netzen. Diss. B, TH Karl-Marx-Stadt 1985.