



TECHNISCHE UNIVERSITÄT CHEMNITZ

## **Sonderforschungsbereich 393**

Parallele Numerische Simulation für Physik und Kontinuumsmechanik

Arnd Meyer

Peter Steinhorst

### **Überlegungen zur Parameterwahl im Bramble-Pasciak-CG für gemischte FEM**

Preprint SFB393/05-07

Preprintreihe des Chemnitzer SFB 393

ISSN 1619-7178 (Print)

ISSN 1619-7186 (Internet)

**SFB393/05-07**

**April 2005**

# Inhaltsverzeichnis

<b>1</b>	<b>Einleitung</b>	<b>1</b>
<b>2</b>	<b>Spektralbetrachtungen der transformierten Systemmatrix</b>	<b>2</b>
2.1	Abschätzung der Spektren . . . . .	2
2.2	Wahl eines geeigneten Parameters . . . . .	3
2.3	Berechnung der Rayleighquotienten . . . . .	6
<b>3</b>	<b>Numerische Experimente</b>	<b>7</b>

Author's addresses:

Arnd Meyer  
TU Chemnitz  
Fakultät für Mathematik  
D-09107 Chemnitz

Peter Steinhorst  
TU Chemnitz  
Fakultät für Mathematik  
D-09107 Chemnitz

email: [peter.steinhorst@mathematik.tu-chemnitz.de](mailto:peter.steinhorst@mathematik.tu-chemnitz.de)

<http://www.tu-chemnitz.de/sfb393/>

# 1 Einleitung

Für die Beschreibung einiger Probleme aus Physik und Kontinuumsmechanik werden gemischte Formulierungen verwendet. Beispiele hierfür sind Strömungsprobleme (Stokes), Elastizitätsprobleme auch inkompressibler Materialien sowie piezoelektrische Probleme. Die numerische Approximation mit gemischter FEM führt dabei häufig auf lineare Gleichungssysteme der folgenden Form:

$$\begin{pmatrix} K & B \\ B^T & -C \end{pmatrix} \begin{pmatrix} u \\ p \end{pmatrix} = \begin{pmatrix} f \\ g \end{pmatrix}. \quad (1)$$

Wir betrachten den Fall mit symmetrisch positiv definiten Matrix  $K \in \mathbb{R}^{N_u \times N_u}$ , die Matrix  $C \in \mathbb{R}^{N_p \times N_p}$  soll symmetrisch positiv semidefinit sein. Die Gesamtsystemmatrix

$$A = \begin{pmatrix} K & B \\ B^T & -C \end{pmatrix} \in \mathbb{R}^{N \times N}$$

ist dann symmetrisch, aber im Allgemeinen indefinit. Dies sieht man leicht bei der Wahl eines „unteren Blockvektors“:

Mit  $\mathbf{x} = \begin{pmatrix} 0 \\ q \end{pmatrix}$  gilt  $(A\mathbf{x}, \mathbf{x}) = -(Cq, q) \leq 0$ .

In [1] beschreiben Bramble und Pasciak eine Methode, das indefinite System (1) mit einer geeigneten Vorkonditionierung zu einem positiv definiten System umzuformen. Dabei wird zunächst angenommen, dass  $K_0$  s.p.d. ein guter Vorkonditionierer für  $K$  sei und das System (1) von links mit der Matrix

$$\begin{pmatrix} K_0^{-1} & O \\ B^T K_0^{-1} & -I \end{pmatrix}$$

multipliziert. Die Systemmatrix des transformierten Systems

$$\begin{pmatrix} K_0^{-1}K & K_0^{-1}B \\ B^T K_0^{-1}(K - K_0) & B^T K_0^{-1}B + C \end{pmatrix} \begin{pmatrix} u \\ p \end{pmatrix} = \begin{pmatrix} K_0^{-1}f \\ B^T K_0^{-1}f - g \end{pmatrix}$$

ist s.p.d. in einem speziell gewählten Skalarprodukt.

Wir betrachten eine von Meyer und Steidten in [2] eingeführte Verallgemeinerung: Eine Matrix  $B_0$  wird als Vorkonditionierer für das Schur-Komplement  $B^T K_0^{-1}B + C$  gewählt, wobei positive Konstanten  $\underline{\beta}, \bar{\beta}$  existieren sollen, so dass

$$\underline{\beta}(B_0q, q) \leq ((B^T K_0^{-1}B + C)q, q) \leq \bar{\beta}(B_0q, q) \quad \forall q \in \mathbb{R}^{N_p}.$$

Ferner werden 2 positive Skalierungsfaktoren  $\delta$  und  $\gamma$  hinzugefügt. Nun wird (1) von links mit der Matrix

$$\mathcal{K} = \begin{pmatrix} K_0^{-1} & O \\ \delta B_0^{-1} B^T K_0^{-1} & -\gamma \delta B_0^{-1} \end{pmatrix}$$

multipliziert. Die neue Systemmatrix  $\mathcal{K}A$  bezeichnen wir mit

$$\mathcal{A} = \begin{pmatrix} K_0^{-1}K & K_0^{-1}B \\ \delta B_0^{-1} B^T (K_0^{-1}K - \gamma I) & \delta B_0^{-1} (B^T K_0^{-1}B + \gamma C) \end{pmatrix}.$$

Wir definieren im folgenden das Schur-Komplement mit

$$S = B^T K_0^{-1} B + \gamma C.$$

$B_0$  soll dabei nun ein guter Vorkonditionierer für  $S$  sein. Wählt man  $\gamma$  so, dass  $K - \gamma K_0$  positiv definit ist, lässt sich zeigen dass  $\mathcal{A}$  bezüglich eines speziellen Skalarproduktes  $\langle \cdot, \cdot \rangle$  symmetrisch positiv definit ist (nachzulesen in [2]). Führt man für die Vektoren die Blockschreibweise

$$\mathbf{x} = \begin{pmatrix} x \\ p \end{pmatrix}, \mathbf{y} = \begin{pmatrix} y \\ q \end{pmatrix}$$

ein, so wird dieses Skalarprodukt definiert durch

$$\langle \mathbf{x}, \mathbf{y} \rangle := ((K - \gamma K_0)x, y) + (\delta^{-1} B_0 p, q) =: \langle x, y \rangle_o + \langle p, q \rangle_u, \quad (2)$$

wobei  $(\cdot, \cdot)$  das übliche euklidische Skalarprodukt im  $\mathbb{R}^{N_u}$  bzw.  $\mathbb{R}^{N_p}$  bezeichnet. Man kann also unter Verwendung von  $\mathcal{A}$  und  $\langle \cdot, \cdot \rangle$  den „handelsüblichen“ CG benutzen um das System zu lösen, und hat quasi nebenbei gleich noch eine (blockweise) Vorkonditionierung mit eingebaut. Bisher noch offen ist eine geeignete Wahl der Parameter  $\delta$  und  $\gamma$ . Im Folgenden soll eine Strategie erörtert werden, einen möglichst günstigen Wert für  $\delta$  zu finden.

## 2 Spektralbetrachtungen der transformierten Systemmatrix

### 2.1 Abschätzung der Spektren

Beim Nachweis der positiven Definitheit von  $\mathcal{A}$  bzgl.  $\langle \cdot, \cdot \rangle$  bewiesen Meyer und Steidten in [2] ein Lemma, welches die Spektraleigenschaften von  $\mathcal{A}$  abschätzt gegenüber dem Blockdiagonalteil von  $\mathcal{A}$ , hier bezeichnet als

$$\mathcal{M} = \begin{pmatrix} K_0^{-1} K & O \\ O & \delta B_0^{-1} S \end{pmatrix}.$$

Es existieren nämlich positive Konstanten  $\underline{\alpha}, \bar{\alpha}$  mit

$$\underline{\alpha} \langle \mathcal{M} \mathbf{x}, \mathbf{x} \rangle \leq \langle \mathcal{A} \mathbf{x}, \mathbf{x} \rangle \leq \bar{\alpha} \langle \mathcal{M} \mathbf{x}, \mathbf{x} \rangle.$$

Das liefert als eine Abschätzung für die Konditionszahl von  $\mathcal{A}$ :

$$\kappa(\mathcal{A}) \leq \frac{\bar{\alpha}}{\underline{\alpha}} \kappa(\mathcal{M}).$$

Ziel ist es damit, die Kondition von  $\mathcal{M}$  möglichst klein zu halten, um die Kondition von  $\mathcal{A}$  zu beschränken.

Für das Spektrum von  $\mathcal{M}$  gilt aufgrund der Blockdiagonalstruktur:

$$\sigma(\mathcal{M}) = \sigma(K_0^{-1} K) \cup \sigma(\delta B_0^{-1} S) = \sigma(K_0^{-1} K) \cup \delta \sigma(B_0^{-1} S).$$

Der Parameter  $\delta$  skaliert also das Spektrum des unteren Blockes gegenüber dem oberen Block. Um im CG eine schnelle Konvergenz zu erreichen möchten wir mittels geschickter Wahl von  $\delta$  nun die beiden Blockspektren möglichst ineinanderschieben, damit die Kondition der Gesamtmatrix möglichst klein bleibt. Fernerhin bezeichnen wir mit  $\lambda_{\min}(T)$  und  $\lambda_{\max}(T)$  den betragsmäßig kleinsten bzw. größten Eigenwert einer Matrix  $T$ . Wir betrachten nun die Blöcke von  $\mathcal{M}$  bezüglich der in (2) definierten Skalarprodukte. Es ist leicht einzusehen, daß  $K_0^{-1}K$  im Skalarprodukt  $\langle \cdot, \cdot \rangle_o$  selbstadjungiert ist, da gilt:

$$\begin{aligned} \langle K_0^{-1}Kx, y \rangle_o &= ((K - \gamma K_0)K_0^{-1}Kx, y) = ((KK_0^{-1}K - \gamma K)x, y) \\ &= (KK_0^{-1}(K - \gamma K_0)x, y) = ((K - \gamma K_0)x, K_0^{-1}Ky) \\ &= \langle x, K_0^{-1}Ky \rangle_o . \end{aligned}$$

Analog ist  $B_0^{-1}S$  wie auch  $\delta B_0^{-1}S$  bezüglich  $\langle \cdot, \cdot \rangle_u$  selbstadjungiert, hier gilt mit der offensichtlichen Symmetrie des Schurkomplementes  $S$  bezüglich  $(\cdot, \cdot)_u$ :

$$\langle \delta B_0^{-1}Sq, q \rangle_u = (Sq, q) = (q, Sq) = (\delta^{-1}B_0q, \delta B_0^{-1}Sq) = \langle q, \delta B_0^{-1}Sq \rangle_u .$$

Bekanntlich gilt somit die folgende Beziehung zwischen den Rayleigh-Quotienten und den Eigenwerten:

$$\lambda_{\min}(K_0^{-1}K) \leq R_o(x) := \frac{\langle K_0^{-1}Kx, x \rangle_o}{\langle x, x \rangle_o} \leq \lambda_{\max}(K_0^{-1}K) \quad (3)$$

$$\lambda_{\min}(B_0^{-1}S) \leq R_u(q) := \frac{\langle B_0^{-1}Sq, q \rangle_u}{\langle q, q \rangle_u} \leq \lambda_{\max}(B_0^{-1}S) \quad (4)$$

Unter der Bedingung, daß  $K_0$  und  $B_0$  gute Vorkonditionierer für die Matrizenblöcke  $K$  bzw.  $S$  sind, sind die Konditionszahlen

$$\kappa(K_0^{-1}K) = \frac{\lambda_{\max}(K_0^{-1}K)}{\lambda_{\min}(K_0^{-1}K)} \quad \text{sowie} \quad \kappa(B_0^{-1}S) = \frac{\lambda_{\max}(B_0^{-1}S)}{\lambda_{\min}(B_0^{-1}S)}$$

im wesentlichen von der Größenordnung  $\mathcal{O}(1)$ . Unmittelbar folgt daraus aber auch für das Verhältnis zwischen Rayleigh-Quotienten prinzipiell beliebiger Vektoren  $x, y \in \mathbb{R}^{N_u} \neq 0$  bzw.  $p, q \in \mathbb{R}^{N_p} \neq 0$ :

$$\frac{R_o(x)}{R_o(y)} \leq \kappa(K_0^{-1}K), \quad \frac{R_u(p)}{R_u(q)} \leq \kappa(B_0^{-1}S),$$

so daß der Quotient zwischen Rayleigh-Quotienten zu verschiedenen nichtverschwindenden Vektoren immer von der Größenordnung  $\mathcal{O}(1)$  ist.

## 2.2 Wahl eines geeigneten Parameters

Die Kondition der Blockdiagonalmatrix  $\mathcal{M}$  ergibt sich aus der Vereinigungsmenge der Eigenwerte von  $K_0^{-1}K$  und  $\delta B_0^{-1}S$ :

$$\kappa(\mathcal{M}) = \frac{\max(\lambda_{\max}(K_0^{-1}K), \delta \cdot \lambda_{\max}(B_0^{-1}S))}{\min(\lambda_{\min}(K_0^{-1}K), \delta \cdot \lambda_{\min}(B_0^{-1}S))} \quad (5)$$

Angenommen, wir würden  $\delta$  so wählen, daß  $\lambda_{\min}(K_0^{-1}K) = \delta\lambda_{\min}(B_0^{-1}S)$  ist. Dann folgt aus obiger Formel sofort

$$\kappa(\mathcal{M}) = \max(\kappa(K_0^{-1}K), \kappa(B_0^{-1}S))$$

Hier werden die Spektren der Blöcke so ineinandergeschoben, daß der untere Rand übereinstimmt. Die gleiche Abschätzung ergibt sich auch, wenn man stattdessen  $\lambda_{\max}(K_0^{-1}K) = \delta\lambda_{\max}(B_0^{-1}S)$  fordert, oder aber

$$\lambda_{\max}(K_0^{-1}K) \left\{ \begin{array}{c} + \\ \cdot \end{array} \right\} \lambda_{\min}(K_0^{-1}K) = \delta\lambda_{\max}(B_0^{-1}S) \left\{ \begin{array}{c} + \\ \cdot \end{array} \right\} \delta\lambda_{\min}(B_0^{-1}S)$$

setzt, was eine Aufeinanderschiebung des oberen Spektrumrandes bzw. der (arithmetischen oder geometrischen) Spektrenmittelpunkte bedeutet. Veranschaulicht werden die beschriebenen Situationen in Abb. 1, wobei  $T_1$  und  $T_2$  die Matrizen  $K_0^{-1}K$  und  $\delta B_0^{-1}S$  derart bezeichnen, daß  $\kappa(T_1) \geq \kappa(T_2)$  gilt. Dabei wird mit  $\overline{\sigma(T_i)}$  die konvexe Hülle von  $\sigma(T_i)$  bezeichnet. Das Problem liegt nun darin, daß wir  $\delta$  nicht einfach nach einer der beschriebenen

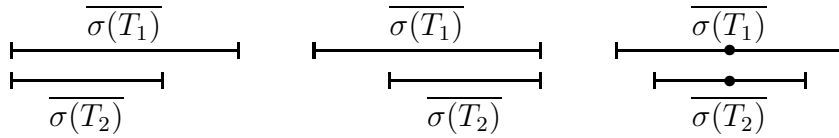


Abbildung 1: Spektrensachtelung bei optimaler Wahl von  $\delta$

Varianten wählen können, da wir die extremalen Eigenwerte von  $K_0^{-1}K$  sowie  $B_0^{-1}S$  von vorneherein nicht kennen. Wir wissen aber, daß beliebig gewählte Rayleigh-Quotienten immer innerhalb der Spektromsgrenzen liegen. Wählen wir einen Vektor in Blockaufteilung  $v = \begin{pmatrix} x \\ q \end{pmatrix}$ ,  $x \in \mathbb{R}^{N_u} \neq 0$ ,  $q \in \mathbb{R}^{N_p} \neq 0$ , und wählen wir  $\delta$  so, daß  $R_o(x) = \delta R_u(q)$  ist, so ergibt sich mittels (3),(4) sofort:

$$R_o(x) \in [\lambda_{\min}(K_0^{-1}K), \lambda_{\max}(K_0^{-1}K)], R_o(x) \in [\lambda_{\min}(\delta B_0^{-1}S), \lambda_{\max}(\delta B_0^{-1}S)].$$

Immerhin haben wir die Tatsache, daß bei guter Vorkonditionierung  $\lambda_{\min}$  und  $\lambda_{\max}$  jeweils „nah beieinanderliegen“, die Wahl von

$$\delta = \frac{R_o(x)}{R_u(q)} \tag{6}$$

führt also schon einmal in die richtige Größenordnung. Wir erhalten mit diesem  $\delta$  nun eine Abschätzung für die Kondition von  $\mathcal{M}$ . In Gleichung (5) können wir dazu eine Fallunterscheidung durchführen:

- $\lambda_{\min}(K_0^{-1}K) \leq \delta \lambda_{\min}(B_0^{-1}S)$

$$\begin{aligned}
\Rightarrow \kappa(\mathcal{M}) &= \max\left(\frac{\lambda_{\max}(K_0^{-1}K)}{\lambda_{\min}(K_0^{-1}K)}, \frac{\delta \lambda_{\max}(B_0^{-1}S)}{\lambda_{\min}(K_0^{-1}K)}\right) \\
&\stackrel{(6)}{=} \max\left(\kappa(K_0^{-1}K), \frac{R_o(x)\lambda_{\max}(B_0^{-1}S)}{R_u(q)\lambda_{\min}(K_0^{-1}K)}\right) \\
&\stackrel{(3),(4)}{\leq} \max\left(\kappa(K_0^{-1}K), \frac{\lambda_{\max}(K_0^{-1}K)\lambda_{\max}(B_0^{-1}S)}{\lambda_{\min}(B_0^{-1}S)\lambda_{\min}(K_0^{-1}K)}\right) \\
&= \max(\kappa(K_0^{-1}K), \kappa(K_0^{-1}K)\kappa(B_0^{-1}S)) \\
&\stackrel{\kappa(\cdot) \geq 1}{=} \kappa(K_0^{-1}K)\kappa(B_0^{-1}S).
\end{aligned}$$

- $\lambda_{\min}(K_0^{-1}K) > \delta \lambda_{\min}(B_0^{-1}S)$

$$\begin{aligned}
\Rightarrow \kappa(\mathcal{M}) &= \max\left(\frac{\lambda_{\max}(K_0^{-1}K)}{\delta \lambda_{\min}(B_0^{-1}S)}, \frac{\delta \lambda_{\max}(B_0^{-1}S)}{\delta \lambda_{\min}(B_0^{-1}S)}\right) \\
&\stackrel{(6)}{=} \max\left(\frac{R_u(q)\lambda_{\max}(K_0^{-1}K)}{R_u(q)R_o(x)\lambda_{\min}(B_0^{-1}S)}, \kappa(B_0^{-1}S)\right) \\
&\stackrel{(3),(4)}{\leq} \max\left(\frac{\lambda_{\max}(B_0^{-1}S)\lambda_{\max}(K_0^{-1}K)}{\lambda_{\min}(K_0^{-1}K)\lambda_{\min}(B_0^{-1}S)}, \kappa(B_0^{-1}S)\right) \\
&= \max(\kappa(K_0^{-1}K)\kappa(B_0^{-1}S), \kappa(B_0^{-1}S)) \\
&= \kappa(K_0^{-1}K)\kappa(B_0^{-1}S).
\end{aligned}$$

Die grafische Veranschaulichung dieses Sachverhaltes liefert Abb. 2.

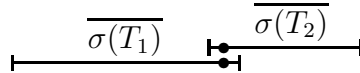


Abbildung 2: Spektrenüberlappung mittels einfacher Wahl von  $\delta$

Lediglich durch Berechnung je eines oberen und unteren Rayleigh-Quotienten läßt sich damit absichern, daß die Kondition von  $\mathcal{M}$  das Produkt der Konditionszahlen der Blöcke nicht übersteigt. Sind die verwendeten Vorkonditionierer hinreichend gut (d.h.  $\kappa(K_0^{-1}K)$ ,  $\kappa(B_0^{-1}S)$  jeweils  $\mathcal{O}(1)$ ), so ist bei Wahl von  $\delta$  vermöge (6) auch  $\kappa(\mathcal{M})$  von der Größenordnung  $\mathcal{O}(1)$ , wobei allerdings die Konstanten im ungünstigsten Fall multipliziert werden. Selbst eine logarithmische Abhängigkeit der Kondition von  $N$  via

$$\kappa(K_0^{-1}K) = \mathcal{O}((\ln(N_u))^k), \quad \kappa(B_0^{-1}S) = \mathcal{O}((\ln(N_u))^l)$$

führt für die Gesamtkondition auf eine lediglich logarithmische Abhängigkeit

$$\kappa(\mathcal{M}) \leq \mathcal{O}((\ln(N))^{k+l}).$$

## 2.3 Berechnung der Rayleighquotienten

Eine Problematik ist nun noch die Berechnung von  $R_o$  bzw.  $R_u$ . Wir haben ja z.B. den Vorkonditionierer  $K_0$  nicht direkt vorliegen, sondern können nur einen gegebenen Vektor mit  $K_0^{-1}$  multiplizieren. Wir können aber über Umformungen berechenbare Ausdrücke erzeugen:

$$\begin{aligned}
R_o &= \frac{\langle K_0^{-1}Kx, x \rangle_o}{\langle x, x \rangle_o} = \frac{\langle (K - \gamma K_0)K_0^{-1}Kx, x \rangle}{\langle (K - \gamma K_0)x, x \rangle} \\
&= \frac{\langle (KK_0^{-1}K - \gamma K)K_0^{-\frac{1}{2}}y, K_0^{-\frac{1}{2}}y \rangle}{\langle (K - \gamma K_0)K_0^{-\frac{1}{2}}y, K_0^{-\frac{1}{2}}y \rangle} \quad \leftarrow x := K_0^{-\frac{1}{2}}y \\
&= \frac{\langle ((K_0^{-\frac{1}{2}}KK_0^{-\frac{1}{2}})^2 - \gamma K_0^{-\frac{1}{2}}KK_0^{-\frac{1}{2}})y, y \rangle}{\langle (K_0^{-\frac{1}{2}}KK_0^{-\frac{1}{2}} - \gamma I)y, y \rangle} \\
&= \frac{\langle \tilde{K}(\tilde{K} - \gamma I)y, y \rangle}{\langle (\tilde{K} - \gamma I)y, y \rangle} \quad \leftarrow \tilde{K} := K_0^{-\frac{1}{2}}KK_0^{-\frac{1}{2}} \\
&= \frac{\langle \tilde{K}(\tilde{K} - \gamma I)^{\frac{1}{2}}y, (\tilde{K} - \gamma I)^{\frac{1}{2}}y \rangle}{\langle (\tilde{K} - \gamma I)^{\frac{1}{2}}y, (\tilde{K} - \gamma I)^{\frac{1}{2}}y \rangle} \quad \leftarrow \tilde{K}, (\tilde{K} - \gamma I)^{\frac{1}{2}} \text{ vertauschbar} \\
&= \frac{\langle \tilde{K}z, z \rangle}{\langle z, z \rangle} = \frac{\langle KK_0^{-\frac{1}{2}}z, K_0^{-\frac{1}{2}}z \rangle}{\langle z, z \rangle} \quad \leftarrow z := (\tilde{K} - \gamma I)^{\frac{1}{2}}y \\
&= \frac{\langle KK_0^{-1}w, K_0^{-1}w \rangle}{\langle K_0^{-1}w, w \rangle} \quad \leftarrow z := K_0^{-\frac{1}{2}}w
\end{aligned}$$

Die Vertauschbarkeit von  $\tilde{K}$  und  $(\tilde{K} - \gamma I)^{\frac{1}{2}} = f(\tilde{K})$  ist klar aufgrund der Eigenzerlegung der Matrixfunktion ( $\tilde{K}$  symmetrisch).

Damit lässt sich also tatsächlich  $R_o$  berechnen, da wir lediglich Vektoren mit  $K$  bzw.  $K_0^{-1}$  multiplizieren müssen:

$$R_o = \frac{\langle Kv, v \rangle}{\langle v, w \rangle} \quad \text{mit } v = K_0^{-1}w. \quad (7)$$

Für den unteren Teil gilt

$$R_u = \frac{\langle B_0^{-1}Sq, q \rangle_u}{\langle q, q \rangle_u} = \frac{\langle \delta^{-1}Sq, q \rangle}{\langle \delta^{-1}B_0q, q \rangle} = \frac{\langle Sq, q \rangle}{\langle B_0q, q \rangle}.$$

Wenn wir  $B_0$  direkt kennen (z.B. Diagonalvorkonditionierung für die Massenmatrix) können wir somit direkt  $R_u$  berechnen, ansonsten folgt mit  $q = B_0^{-1}r$  die Äquivalenz

$$R_u = \frac{\langle Sq, q \rangle}{\langle B_0q, q \rangle} = \frac{\langle SB_0^{-1}r, B_0^{-1}r \rangle}{\langle B_0^{-1}r, r \rangle}. \quad (8)$$



**Ideen für optimalere Parameterwahl:** Es stellt sich nun die Frage ob es möglich ist, sich dem Optimalzustand aufeinanderliegender Spektrenmittelpunkte (vgl. Abb. 1) noch weiter anzunähern. Man könnte z.B. für jeweils mehrere Vektoren, die im Idealfall nahe an Eigenvektoren relativ großer und relativ kleiner Eigenwerte liegen, Rayleighquotienten berechnen und aus dem größten und kleinsten das arithmetische oder geometrische (dieses wird aber sehr stark verfälscht falls Ausreißer nahe 0 auftreten) Mittel bilden. Mit rein „zufälliger“ Vektorenwahl ohne weiteres Wissen über die Eigenwerte/-vektoren ist allerdings kein qualitativ besseres Ergebnis zu erwarten. Es wären dann die Kosten für die Berechnung von Rayleigh-Quotienten gegen die von ggf. einzusparenden Iterationsschritten im CG abzuwägen.

### 3 Numerische Experimente

Erste Experimente wurden mit dem zweidimensionalen Elastizitätsprogramm *SPC-Pm2AdMix* durchgeführt, um an einigen Beispielen die Auswirkung der Parameterwahl zu testen. Als erstes Beispiel wurde hier ein Quadrat aus normal elastischem Material (z.B. Stahl) genommen, in welchem ein kleineres Quadrat aus einem inkompressiblen Material eingelassen ist. Unten ist ein Dirichletrand vorgegeben, während oben auf den mittleren Teil eine konstante Kraft wirkt (siehe dazu auch Abb. 3).

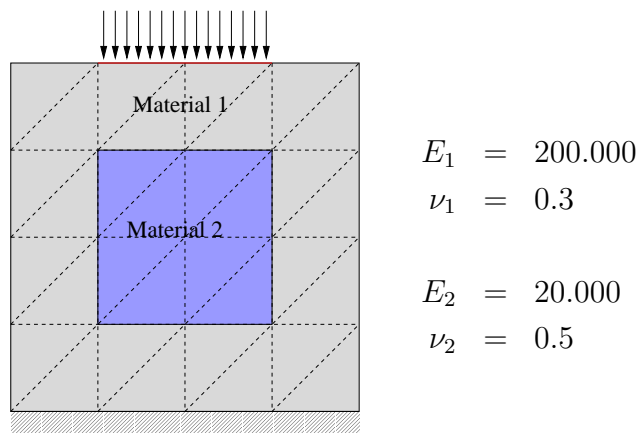


Abbildung 3: Grafische Veranschaulichung von Testbeispiel 1

Zunächst einmal wurde ausgehend vom Grobnetz (Level 0) für  $\delta$ -Werte aus verschiedenen Größenordnungen eine Rechnung mit 4-facher Totalverfeinerung durchgeführt, um die dabei auftretenden MixCG-Iterationszahlen in jedem Level sowie die berechneten Maximalwerte der Verschiebung  $u_x$ ,  $u_y$  sowie des hydrostatischen Druckes  $p$  im feinsten Level 4 zu vergleichen. Zuvor wurde noch mittels Testrechnungen ein möglichst großer geeigneter Wert für  $\gamma$  bestimmt, so daß eine Verletzung der positiven Definitheit von  $K - \gamma K_0$  nicht feststellbar war. Für Dreieckselemente mit globaler roter Verfeinerung ergab sich hierbei ein Wert von  $\gamma = 0.005$  sowie die in Abbildung 4 dargestellte Abhängigkeit von  $\delta$ .

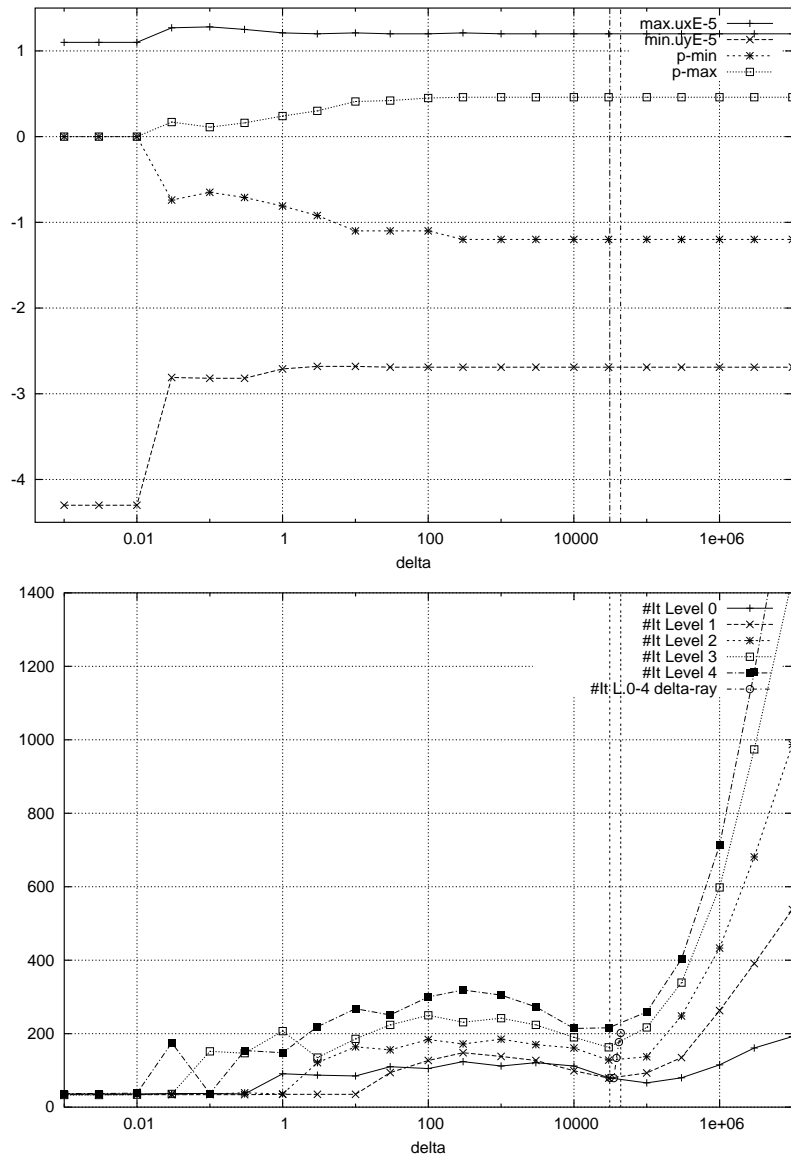


Abbildung 4: Lösungskomponenten und Iterationszahlen in Bsp.1 bei Dreiecksvernetzung

Die automatisch vermöge (6) gewählten Werte  $\delta =: \delta_{ray}$  liegen abhängig vom Level etwa im Bereich 30000..45000, im Diagramm angedeutet durch senkrechte Linien. Als Ausgangsvektoren für die Berechnung mittels (7),(8) wurden dabei Vektoren aus lauter Einsen gewählt. Das Diagramm zeigt kleine Iterationszahlen (jeweils zwischen 30 und 40 Iterationen in allen Leveln) für kleine  $\delta \leq 0.01$ , danach ein in den einzelnen Leveln etwas chaotisches Einsetzen größerer Iterationszahlen für etwa  $0.01 < \delta \leq 10$  sowie anscheinend stetig von  $\delta$  abhängige Iterationszahlen für etwa  $\delta > 10$ . Die Kurven der Iterationszahlen zeigen hier ein relatives Minimum etwa in Höhe des Bereichs der automatisch ermittelten  $\delta_{ray}$ ,

während ein kleinerer „Buckel“ bei  $\delta$ -Werten zwischen 100 und 1000 sowie sehr schnell ansteigende Iterationszahlen für  $\delta$  oberhalb etwa 100000 zu beobachten sind. Die berechneten Lösungskomponenten zeigen ein stabiles Verhalten mit etwa gleichbleibenden Werten für  $\delta$  im Bereich um  $\delta_{\text{ray}}$  herum, während abweichende  $p$ -Werte schon bei etwa  $\delta < 100$  sowie signifikant abweichende Verschiebungen bei etwa  $\delta \leq 1$  ermittelt werden. Etwa analog sind die Ergebnisse der gleichen Testrechnungen bei Vierecksvernetzung, welche in Abbildung 5 dargestellt sind.

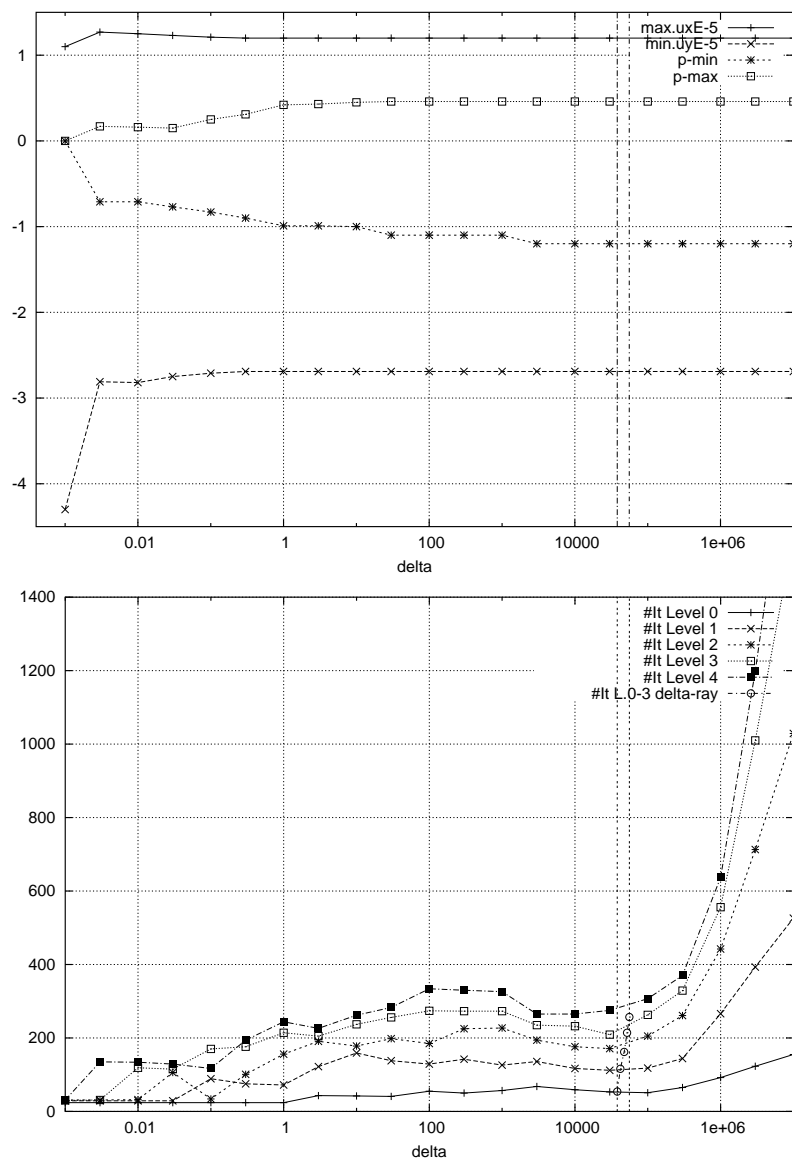


Abbildung 5: Lösungskomponenten und Iterationszahlen in Bsp.1 bei Vierecksvernetzung

Hier wurde ein ebenfalls durch „Probieren“ ermittelter Wert von  $\gamma = 0.004$  verwendet.

Beim genaueren Vergleich der Vernetzungsarten ist ersichtlich, daß die Abweichung zumindest der Verschiebungskomponenten der Lösung bei Verwendung von Viereckselementen erst bei um etwa um eine Größenordnung kleinerem  $\delta$  eintritt, obwohl das Netz durch Zusammenlegen zweier Dreiecke zu einem Viereck etwas weniger Freiheitsgrade besitzt. Grund dafür ist vermutlich, daß die gesamte Anordnung symmetrisch ist und die Vierecksvernetzung im Gegensatz zur Dreiecksvernetzung ebenfalls diese Symmetrie aufweist.

Testweise wurde auf einer „schnellen“ Maschine einmal mit adaptiver Verfeinerung bis zur Überschreitung von 10000 Knoten gerechnet. Die dabei aufgelaufenen PCG-Zeiten, sowie die prozentuale Abweichung der maximalen Verschiebungskomponenten  $u_x, u_y$  vom vermutlich korrekten Wert (näherungsweise stabil berechnet für Werte von  $\delta \geq 1$ ) verdeutlicht Abbildung 6.

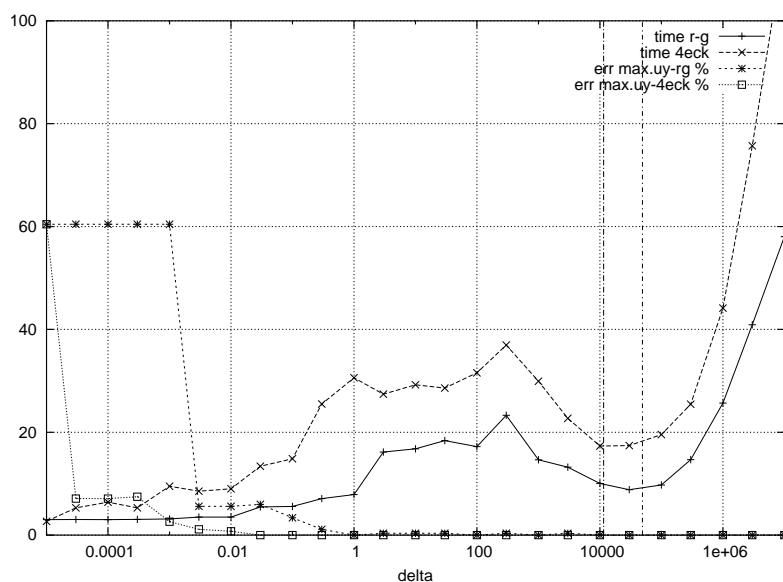


Abbildung 6: PCG-Zeiten für adaptive Rechnung in Bsp.1 bei rot-grüner Dreiecksvernetzung sowie Vierecksvernetzung mit max. Anisotropiefaktor 2.1

Der Effekt abnehmender Iterationszeiten und signifikant wachsendem Fehler bei  $\delta < 1$  ist darauf zurückzuführen, daß sich bei sehr kleinem (und evtl. auch sehr großem, was hier nicht mitgerechnet wurde)  $\delta$  die Blöcke der Gesamtmatrix  $\mathcal{A}$  in der Größenordnung wesentlich unterscheiden. Der gering gewichtete Block wird also in der Rechnung mit wachsendem Ungleichgewicht immer stärker vernachlässigt, und spätestens bei einem Größenordnungsunterschied von mehr als  $\varepsilon^2$ , wobei  $\varepsilon$  das Abbruch-Epsilon darstellt, erzeugt der PCG keine Konvergenz mehr gegen eine sinnvolle Lösung. Hier greift dann nur noch der obere bzw. untere Block, wo durch die jeweils gute Vorkonditionierung eine sehr schnelle Konvergenz eintritt. Bei automatischer Wahl von  $\delta = \delta_{\text{ray}}$  (ähnlich für  $\delta$ -Werte in der gleichen Größenordnung) ist in diesem Beispiel erfreulicherweise nur ein moderates Wachstum der Iterationszahlen festzustellen: Bei adaptiver rot-grüner Verfeinerung 80 Iterationen beim

Start mit 81 Knoten, 155 It. bei 1169 Knoten, 232 It. bei 10838 Knoten. Etwa vergleichbar ist die Situation bei totaler Verfeinerung: 177 Iterationen in Level 3 (4225 Knoten), 202 Iterationen in Level 4 (16641 Knoten).

Als ein weiteres Beispiel wurde dieselbe Versuchsanordnung genommen, diesmal aber mit einem Quadrat aus einheitlichem Material der Parameter  $E = 20000$ ,  $\nu = 0.3$ . Einige charakteristische Lösungswerte zeigt Abbildung 7, die Lösungen sind dabei nach 5maliger Totalverfeinerung untersucht worden.

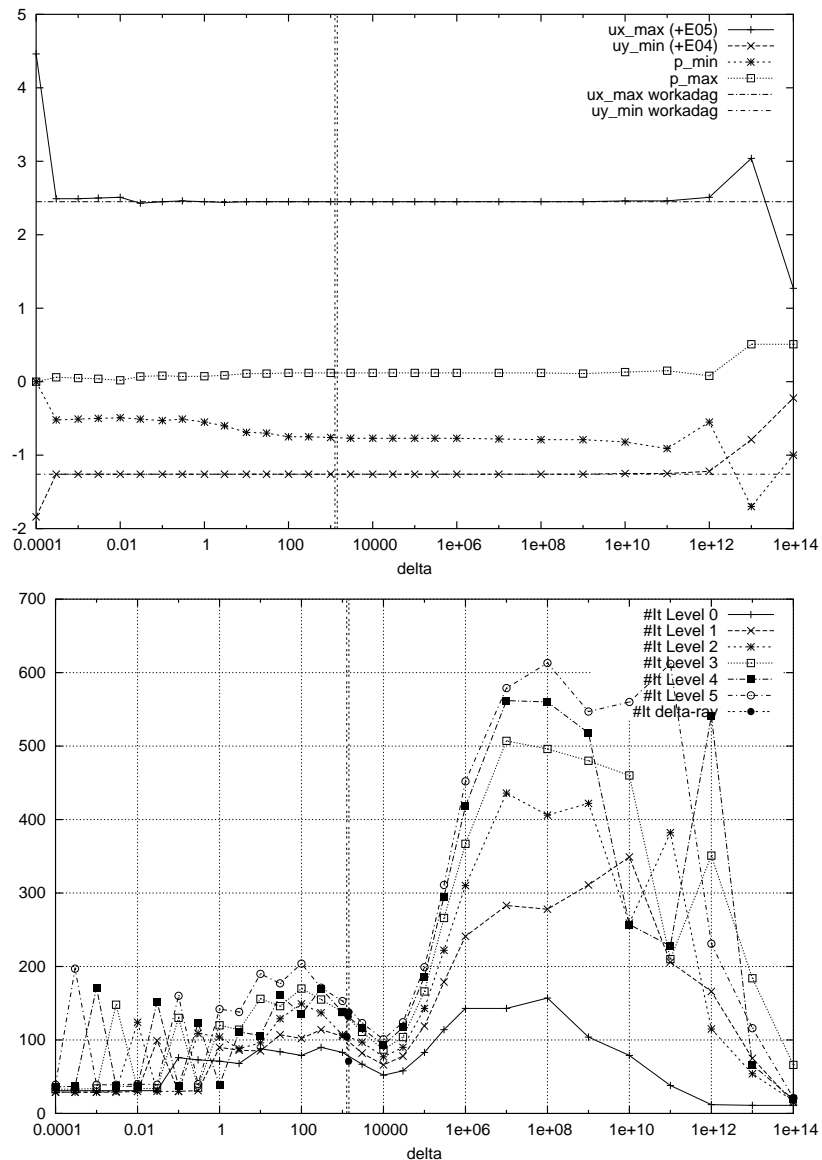


Abbildung 7: Berechnete Lösungskomponenten und Iterationszahlen im Bsp.2

Der Vorteil bei diesem einfacheren Beispiel ist, daß die berechneten Verschiebungskompo-

nenten der Lösung problemlos auch mit dem herkömmlichen Elastizitätsprogramm *SPC-Pm2Ad* berechnet werden können (in Abbildung 7 mit dargestellt) und so ein Vergleich hinsichtlich Korrektheit zumindest der  $u$ -Komponenten möglich ist. Es wurde hier wieder mit global roter Dreiecksverfeinerung gearbeitet und Werte von  $\delta$  über weite Größenordnungen gewählt. Wie oben beschrieben wurde für das gesamte Beispiel ein näherungsweise optimaler Wert von  $\gamma = 0.008$  bestimmt. Auffällig ist, daß für exotische Werte von  $\delta$  außerhalb etwa der Größenordnungen  $10^{-4} \dots 10^{12}$  die Verschiebung  $u_x$  bzw.  $u_y$  keine sinnvollen Werte mehr ergibt. Man beachte, daß dies bei dem verwendeten Abbruch-Epsilon von  $\varepsilon = 10^{-4}$  gerade die von einem „gutartigen“  $\delta$ -Wert  $10^4$  in der Größenordnung um mehr als  $\varepsilon^2$  abweichenden  $\delta$ -Werte sind. Der Bereich stabiler  $p$ -Werte ist dagegen kleiner, der vermutlich korrekte Druck wird zumindest näherungsweise zuverlässig nur für  $\delta \in [100, 10^9]$  getroffen. Die über Rayleigh-Quotienten ermittelten Werte für  $\delta$  lagen hier im Bereich 1250..1500, ein optimales Konvergenz- bzw. Lösungsverhalten scheint aber hier erst bei etwas größerem  $\delta \approx 10^4$  aufzutreten, was auch die Darstellung der Iterationszahlen in Abbildung 7 belegt. Es wäre hier daher zu untersuchen, ob eine zufällige Auswahl weniger Testvektoren noch etwas günstigere Werte bei der automatischen Wahl von  $\delta$  ergibt.

## Literatur

- [1] J.H. Bramble and J.E. Pasciak: A Preconditioning Technique for Indefinite Systems resulting from Mixed Approximations of Elliptic Problems. *Math. Comput.*, 50(181):1-17, 1988.
- [2] A. Meyer and T. Steidten: Improvements and Experiments on the Bramble-Pasciak Type CG for Mixed Problems in Elasticity. Preprint SFB393/01-13, TU Chemnitz 2001.

Other titles in the SFB393 series:

- 03-01 E. Creusé, G. Kunert, S. Nicaise. A posteriori error estimation for the Stokes problem: Anisotropic and isotropic discretizations. January 2003.
- 03-02 S. I. Solov'ëv. Existence of the guided modes of an optical fiber. January 2003.
- 03-03 S. Beuchler. Wavelet preconditioners for the p-version of the FEM. February 2003.
- 03-04 S. Beuchler. Fast solvers for degenerated problems. February 2003.
- 03-05 A. Meyer. Stable calculation of the Jacobians for curved triangles. February 2003.
- 03-06 S. I. Solov'ëv. Eigenvibrations of a plate with elastically attached load. February 2003.
- 03-07 H. Harbrecht, R. Schneider. Wavelet based fast solution of boundary integral equations. February 2003.
- 03-08 S. I. Solov'ëv. Preconditioned iterative methods for monotone nonlinear eigenvalue problems. March 2003.
- 03-09 Th. Apel, N. Düvelmeyer. Transformation of hexahedral finite element meshes into tetrahedral meshes according to quality criteria. May 2003.
- 03-10 H. Harbrecht, R. Schneider. Biorthogonal wavelet bases for the boundary element method. April 2003.
- 03-11 T. Zhanlav. Some choices of moments of refinable function and applications. June 2003.
- 03-12 S. Beuchler. A Dirichlet-Dirichlet DD-pre-conditioner for p-FEM. June 2003.
- 03-13 Th. Apel, C. Pester. Clément-type interpolation on spherical domains - interpolation error estimates and application to a posteriori error estimation. July 2003.
- 03-14 S. Beuchler. Multi-level solver for degenerated problems with applications to p-version of the fem. (*Dissertation*) July 2003.
- 03-15 Th. Apel, S. Nicaise. The inf-sup condition for the Bernardi-Fortin-Raugel element on anisotropic meshes. September 2003.
- 03-16 G. Kunert, Z. Mghazli, S. Nicaise. A posteriori error estimation for a finite volume discretization on anisotropic meshes. September 2003.
- 03-17 B. Heinrich, K. Pönitz. Nitsche type mortaring for singularly perturbed reaction-diffusion problems. October 2003.
- 03-18 S. I. Solov'ëv. Vibrations of plates with masses. November 2003.
- 03-19 S. I. Solov'ëv. Preconditioned iterative methods for a class of nonlinear eigenvalue problems. November 2003.
- 03-20 M. Randrianarivony, G. Brunnett, R. Schneider. Tessellation and parametrization of trimmed surfaces. December 2003.

- 04-01 A. Meyer, F. Rabold, M. Scherzer. Efficient Finite Element Simulation of Crack Propagation. February 2004.
- 04-02 S. Grosman. The robustness of the hierarchical a posteriori error estimator for reaction-diffusion equation on anisotropic meshes. March 2004.
- 04-03 A. Bucher, A. Meyer, U.-J. Görke, R. Kreißig. Entwicklung von adaptiven Algorithmen für nichtlineare FEM. April 2004.
- 04-04 A. Meyer, R. Unger. Projection methods for contact problems in elasticity. April 2004.
- 04-05 T. Eibner, J. M. Melenk. A local error analysis of the boundary concentrated FEM. May 2004.
- 04-06 H. Harbrecht, U. Kähler, R. Schneider. Wavelet Galerkin BEM on unstructured meshes. May 2004.
- 04-07 M. Randrianarivony, G. Brunnett. Necessary and sufficient conditions for the regularity of a planar Coons map. May 2004.
- 04-08 P. Benner, E. S. Quintana-Ortí, G. Quintana-Ortí. Solving Linear Matrix Equations via Rational Iterative Schemes. October 2004.
- 04-09 C. Pester. Hamiltonian eigenvalue symmetry for quadratic operator eigenvalue problems. October 2004.
- 04-10 T. Eibner, J. M. Melenk. An adaptive strategy for hp-FEM based on testing for analyticity. November 2004.
- 04-11 B. Heinrich, B. Jung. The Fourier-finite-element method with Nitsche-mortaring. November 2004.
- 04-12 A. Meyer, C. Pester. The Laplace and the linear elasticity problems near polyhedral corners and associated eigenvalue problems. December 2004.
- 04-13 M. Jung, T. D. Todorov. On the Convergence Factor in Multilevel Methods for Solving 3D Elasticity Problems. December 2004.
- 05-01 C. Pester. A residual a posteriori error estimator for the eigenvalue problem for the Laplace-Beltrami operator. January 2005.
- 05-02 J. Badía, P. Benner, R. Mayo, E. Quintana-Ortí, G. Quintana-Ortí, J. Saak. Parallel Order Reduction via Balanced Truncation for Optimal Cooling of Steel Profiles. February 2005.
- 05-03 C. Pester. CoCoS – Computation of Corner Singularities. April 2005.
- 05-04 A. Meyer, P. Nestler. Mindlin-Reissner-Platte: Einige Elemente, Fehlerschätzer und Ergebnisse. April 2005.
- 05-05 P. Benner, J. Saak. Linear-Quadratic Regulator Design for Optimal Cooling of Steel Profiles. April 2005.

The complete list of current and former preprints is available via  
<http://www.tu-chemnitz.de/sfb393/preprints.html>.