Parallelisierung und Vorkonditionierung des CG-Verfahrens durch Gebietszerlegung

Gundolf Haase und Ulrich Langer
TU Chemnitz, FB Mathematik
Postfach 964, D-O-9010 Chemnitz

Arnd Meyer
Institut für Mechanik
Postfach 408, D-O-9010 Chemnitz

in *Parallele Algorithmen auf Transputersystemen, Teubner-Scripten zur Numerik III, 1992*

Zusammenfassung

Gebietszerlegungstechniken eröffnen eine Vielzahl von Möglichkeiten zur Entwicklung paralleler Lösungsstrategien für Finite-Elemente (F.E.) Gleichungen großer Dimension. Die auf der nichtüberlappenden Gebietszerlegung beruhende F.E. Substrukturtechnik und die damit verbundene Datenverteilung führt zu einer Parallelisierung des CG-Verfahrens, die sich besonders für Mehrprozessorrechner mit lokalem Speicher und Botschaftenaustausch über Links eignet. Für diese parallelisierte Version des CG-Verfahrens werden im vorliegenden Beitrag vier verschiedene Vorkonditionierungen vorgeschlagen, mit denen die Konvergenz des CG-Verfahrens erheblich beschleunigt werden kann, ohne daß im Vergleich zur nichtvorkonditionierten Variante zusätzliche Kommunikation entsteht. Die auf einem Transputerhypercube durchgeführten numerischen Experimente zeigen die Effektivität der vorgeschlagenen Vorkonditionierungen.

Key words: elliptic problems, finite elements, substructuring, domain decomposition, Schwarz' methods, preconditioners, parallel algorithms

AMS (MOS) subject classification: 65N55, 65N22, 65F10, 65Y05, 65Y10, 65F35

1 Einleitung

In den letzten Jahren hat die Parallelrechentechnik eine rasante Entwicklung durchlaufen. Mehrprozessorrechner mit einer beträchtlichen Anzahl an Prozessoren sind z.B. in Form von Transputersystemen kostengünstig auf dem Markt zu erwerben. Im krassen Gegensatz dazu steht die Entwicklung der Software insbesondere der Applikationssoftware. Die Schaffung paralleler Algorithmen in der Finite-Elemente (F.E.) Simulation ist dabei sicherlich ein Schwerpunkt in der angewandten mathematischen Forschung. Gebietszerlegungstechniken eröffnen hierbei eine Vielzahl von Möglichkeiten zur Parallelisierung von F.E. Software. Die nichtüberlappende Gebietszerlegung ist in der FEM unter dem Namen F.E. Substrukturtechnik wohlbekannt und wurde seit Anfang der 60-er Jahre zur effizienten

Erstellung von F.E. Modellen und zur Lösung des F.E. Gleichungssystems auf Rechnern mit beschränktem Hauptspeicher genutzt (siehe z.B. [Prz63]).

Aber bereits 100 Jahre vor dem Einzug der F.E. Substrukturtechnik in die Ingenieurwissenschaften benutzte H. Schwarz eine überlappende Gebietsdekomposition zum konstruktiven Beweis der Existenz harmonischer Funktionen in nichtglatt berandeten Gebieten Ω, die sich als Vereinigung zweier einfacherer, sich überlappender Gebiete Ω_1 und Ω_2, für die die Existenz bekannt ist, darstellen lassen [Sch90]. In dem heute nach Schwarz benannten Iterationsverfahren wird das Randwertproblem (RWP) alternierend in Ω_1 und Ω_2 gelöst, wobei die Randwerte jeweils von der vorhergehenden Iterierten abgenommen werden. S.L. Sobolev gab in [Sob36] die Variationsformulierung der alternierenden Schwarzschen Methode und eröffnete damit die Möglichkeit zu einer allgemeinen Analyse dieses Verfahrens. Der Iterationsoperator der alternierenden Schwarzschen Methode läßt sich als Produkt von Orthoprojektoren darstellen. Wir nennen deshalb das alternierende Schwarzsche Iterationsverfahren im weitern Multiplikative Schwarzsche Methode (MSM). Umfangreiche Untersuchungen zur MSM findet der Leser in den Arbeiten von P.L. Lions [Lio90].

Jeder Zerlegung $\mathbf{V} = \mathbf{V}_1 + \ldots + \mathbf{V}_p$ des energetischen Raums $\mathbf{V} = \mathbf{V}_0$ oder des F.E. Unterraums $\mathbf{V} = \mathbf{V}_h$ in p Teilräume läßt sich durch das Iterationsschema

$$u^{k+1} = (I - \tau P)\, u^k + \tau P u\,; \quad k = 0, 1, \ldots\,; \quad u^0 \in \mathbf{V} \tag{1}$$

eine MSM $(\, P = (I - P_p) \ldots (I - P_1)\,)$ oder eine ASM $(P = P_1 + \ldots + P_p)$ zuordnen, wobei τ ein geeignet gewählter Iterationsparameter ist und $P_i : \mathbf{V} \longrightarrow \mathbf{V}_i$ die Orthoprojektoren in $\mathbf{V}_i$ bzgl. des energetischen Skalarproduktes bezeichnen. Die Additive Schwarzsche Methode (ASM) wurde von M. Dryja und O. Widlund mit Blick auf eine prinzipielle Verbesserung der Parallelisierbarkeit von Schwarzschen Methoden zur Lösung von F.E. Gleichungen vorgeschlagen [DW87]. In Abhängigkeit von der Wahl der Zerlegung des F.E. Raumes $\mathbf{V} = \mathbf{V}_h$ ergeben sich somit eine Vielzahl von Möglichkeiten zur Konstruktion von ASM (siehe [BMS91, DW90, DW91, Xu90] und die dort zitierten Literaturangaben). In der Praxis verwendet man jedoch ausschließlich das CG-Verfahren zur Beschleunigung der ASM als Verfahren der einfachen Iteration. In diesem Sinne betrachten wir die ASM wie auch die MSM als Methoden zur Konstruktion und Analyse von Vorkonditionierungen für das CG-Verfahren als parallelen iterativen Solver von F.E. Gleichungen. Eine Vielzahl von interessanten Arbeiten zu dieser Thematik findet man in den Sammelbänden [CGPW89, CGPW90, GGMP88, Hac91].

Im vorliegenden Beitrag basieren die Ergebnisse zur MSM,ASM und den dazugehörigen Vorkonditionierungsoperatoren auf einer Zerlegung des Gebiets Ω in p sich <u>nicht</u> überlappende Teilgebiete $\Omega_1, \ldots, \Omega_p$ und einer damit verbundenen Zerlegung des F.E. Raumes $\mathbf{V}$ in die Teilräume $\widetilde{\mathbf{V}}_C$ der näherungsweise diskret harmonischen F.E. Funktionen und $\widetilde{\mathbf{V}}_I = \mathbf{V}_{I,1} \oplus \ldots \oplus \mathbf{V}_{I,p}$ aller F.E. Funktionen, die auf den Koppelrändern $\Gamma_C = \bigcup\limits_{i=1}^{p} \partial\Omega_i \setminus \Gamma_D$ verschwinden, wobei Γ_D der Dirichlet-Teil des Randes $\Gamma = \partial\Omega$ ist. Diese Zerlegung wurde von den Autoren in [HLM90] eingeführt und steht in engen Zusammenhang mit der klassischen F.E. Substrukturtechnik. Im Abschnitt 2 der vorliegenden Arbeit betrachten wir deshalb die klassische F.E. Substrukturtechnik und verbinden sie mit der diskret harmonischen Basis. Auf der Grundlage einer mit der klassischen F.E. Substrukturtechnik verbundenen Datenverteilung wird im Abschnitt 3 eine parallelisierte Version des CG-Verfahrens vorgestellt. In der nichtvorkonditionierten Variante, die auf K.H. Law [Law89] zurückgeht, wird eine Kommunikation zwischen den Prozessoren nur zur vollständigen Berechnung der Skalarprodukte und zu einem lokalen Datenaustausch über die Koppelränder hinweg benötigt. Der zuletzt genannte Datenaustausch betrifft nur Daten der Koppelränder und ist folglich im Vergleich zum Gesamtdatenvolumen eine Ordnung kleiner. Von einem Vorkonditionierungsoperator C werden wir deshalb unter anderem verlangen, das der Kommunikationsaufwand gegenüber dem nichtvorkonditionierten Fall $C = I$ nicht erhöht wird. Vorkonditionierungsoperatoren, die dieser Forderung entsprechen stellen wir im Abschnitt 4 vor: Die einfache DD[1]-Vorkonditionierung (Abschnitt 4.1), die ASM-DD-Vorkonditionierung (Abschnitt 4.2), die MSM-DD-Vorkonditionierung (Abschnitt 4.3) und die hierarchische DD-Vorkonditionierung (Abschnitt 4.4). Die hierarchische DD-Vorkonditionierung ist algebraisch äquivalent zur hierachischen Vorkonditionierung von H. Yserentant [Yse86b] und nutzt die Gebietszerlegung nur zur Datenaufteilung. Im abschließenden 5. Abschnitt werden einige Anwendungen auf elliptische RWA für partielle Differentialgleichungen (Dgl.) zweiter Ordnung gegeben und einige Resultate von numerischen Experimenten auf einem Transputerhypercube dargestellt.

2 Die Finite-Elemente-Substrukturtechnik und die diskrete harmonische Basis

Wir betrachten zunächst ein abstraktes, symmetrisches, $\mathbf{V}_0$-elliptisches und $\mathbf{V}_0$-beschränktes Variationsproblem der Form (2),

$$Finde \quad u \in \mathbf{V}_0 : \quad a(u,v) \; = < F, v > \quad \forall v \in \mathbf{V}_0 \quad , \tag{2}$$

welches bei der schwachen Formulierung eines symmetrischen, gleichmäßig elliptischen RWP für eine partielle Dgl. zweiter Ordnung oder für ein System derartiger Dgl. in einem d-dimensionalen (d = 2,3), beschränkten Gebiet Ω mit stückweise glattem Rand $\Gamma = \partial\Omega$ entsteht. Wir zerlegen nun Ω in p sich nichtüberlappende Teilgebiete $\Omega_i \quad (i = 1, 2, \ldots, p)$, die wir im weiteren auch Substrukturen oder Superelemente nennen wollen, so daß $\overline{\Omega} = \bigcup_{i=1}^{p} \overline{\Omega}_i$ und $\Omega_i \cap \Omega_j = \emptyset$ für $i \neq j$. Jedes Teilgebiet $\overline{\Omega}_i$ wird nun seinerseits in finite Elemente $\overline{\delta}_r$ unterteilt, so daß die gesamte Diskretisierung von $\overline{\Omega}_i$ eine reguläre, konforme Triangularisierung ergibt [Cia78]. Im weiteren bezeichnen die Indizes "C" und "I" jeweils Größen, die zu den Knoten des Koppelrandes $\Gamma_C = \bigcup_{i=1}^{p} \partial\Omega_i \setminus \Gamma_D$ bzw. zu den inneren Knoten der Teilgebiete Ω_1,...,Ω_p gehören. Hierbei ist Γ_D immer der Dirichlet-Rand mit entsprechenden Modifikationen im Falle von Differentialgleichungssysteme.

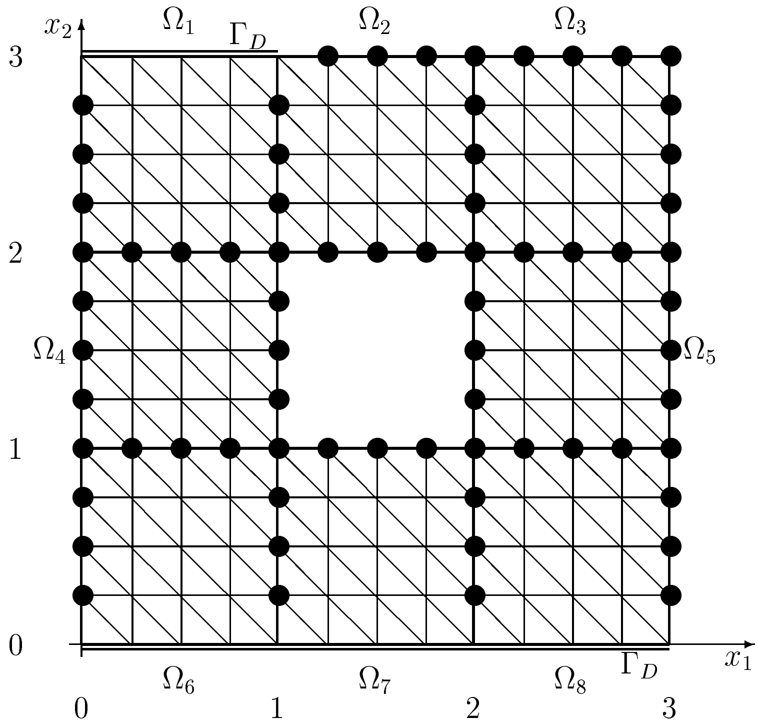
Zur Illustration der Gebietszerlegung betrachten wir ein "ringförmiges" Gebiet Ω (siehe Bild 1) und zerlegen es in die 8 Einheitsquadrate Ω_1 bis Ω_8 . Wie im Bild 1 gezeigt wird, unterteilen wir nun jedes Einheitsquadrat gleichmäßig in rechteckige, dreiknotige (lineare) Dreieckselemente. Im Abschnitt 5.2 dienen uns die Poisson-Gleichung und das lineare Elastizitätsproblem (ebener Verzerrungszustand) im ringförmigen Gebiet Ω mit wesentlichen Randbedingungen auf Γ_D als Modellbeispiel für numerische Experimente.

Es sei

$$\Phi = \left\{ \varphi_1, \cdots, \varphi_{N_C}, \varphi_{N_C+1}, \cdots, \varphi_{N_C+N_{I,1}}, \cdots, \varphi_{N=N_C+N_I} \right\} \quad , \tag{3}$$

die in der FEM übliche Knotenbasis, wobei die ersten N_C Basisfunktionen zu den Knoten des Koppelrandes Γ_C, die nächsten $N_{I,1}$ Basisfunktionen zu den inneren Knoten von Ω_1,

[1]DD = Domain Decomposition (Gebietszerlegung)



$\{\cdot\} \mathrel{\hat{=}}$ "I"

$\{\bullet\} \mathrel{\hat{=}}$ "C"

$$\overline{\Omega} = \bigcup_{i=1}^{8} \overline{\Omega}_i$$

$$\begin{aligned} \Gamma_D = \ & \{(x,0) : \ 0 \le x \le 3\} \\ & \cup \ \{(x,3) : \ 0 \le x \le 1\} \end{aligned}$$

$$\Gamma_C = \bigcup_{i=1}^{8} \partial\overline{\Omega}_i \setminus \Gamma_D$$

Abbildung 1: Gebietszerlegung und Verenetzung (Level 2) des ringförmigen Gebietes Ω.

die nächsten $N_{I,2}$ zu Ω_2 u.s.w. gehören. Folglich gilt: $N_I = N_{I,1} + \cdots + N_{I,p}$. Der F.E. Unterraum

$$\mathbf{V} = \mathbf{V}_h = span(\Phi) = span(\Phi V) = \mathbf{V}_C + \mathbf{V}_I \subset \mathbf{V}_0 \tag{4}$$

kann offenbar als direkte Summe der Unterräume $\mathbf{V}_C = span(\Phi V_C)$ und $\mathbf{V}_I = span(\Phi V_I)$ dargestellt werden, wobei

$$V = (V_C \; V_I) = I = \begin{pmatrix} I_C & O \\ O & I_I \end{pmatrix}_{N \times N} \quad , V_C = \begin{pmatrix} I_C \\ O \end{pmatrix}_{N \times N_C} \quad \text{und} \quad V_I = \begin{pmatrix} O \\ I_I \end{pmatrix}_{N \times N_I}$$

und h den üblichen Diskretisierungsparameter bezeichnet. Der F.E. Isomorphismus zwischen einer F.E. Funktion $u \in \mathbf{V}$ und dem dazugehörigen Vektor $\underline{u} = \left(\underline{u}_C^T, \underline{u}_I^T\right)^T \in \mathbf{R}^N$ der Knotenparameter ist durch die Zuordnung

$$\mathbf{V} \ni u = \Phi V \underline{u} = \Phi \underline{u} \quad \stackrel{\Phi}{\longleftrightarrow} \quad \underline{u} \in \mathbf{R}^N \quad , \tag{5}$$

gegeben.

Die F.E. Approximation

$$\begin{gathered} \text{Finde} \quad u = \Phi V \underline{u} \in \mathbf{V} \ : \\ a(\Phi V \underline{u}, \Phi V \underline{v}) = < F, \Phi V \underline{u} > \ \forall v = \Phi V \underline{v} \in \mathbf{V} \end{gathered} \tag{6}$$

der Variationsgleichung (2) führt bei einmal gewählter Basis Φ von $\mathbf{V}$ unmittelbar auf das System

$$K \, \underline{u} \ = \ \underline{f} \tag{7}$$

der F.E. Gleichungen zur Bestimmung des unbekannten Vektors $\underline{u} \in \mathbf{R}^N$ der Knotenparameter. Die Steifigkeitsmatrix K und der Lastvektor $\underline{f}$ werden entsprechend durch die Identitäten

$$(K\underline{u}, \underline{v}) = (V^T K V \underline{u}, \underline{u}) = a(\Phi V \underline{u}, \Phi V \underline{v}) \quad \forall \underline{u}, \underline{v} \in \mathbf{R}^N \tag{8}$$

und

$$(\underline{f}, \underline{v}) = (\underline{f}, V\underline{u}) = < F, \Phi V \underline{v} > \quad \forall \underline{v} \in \mathbf{R}^N , \tag{9}$$

definiert. Hier und im weiteren bezeichnet $(\cdot, \cdot)$ das übliche Euklidische Skalarprodukt. Die Anordnung der Basisfunktionen (3) hat zur Folge, daß das Gleichungssystem (7) in der Blockform

$$\begin{pmatrix} K_C & K_{CI} \\ K_{IC} & K_I \end{pmatrix} \begin{pmatrix} \underline{u}_C \\ \underline{u}_I \end{pmatrix} = \begin{pmatrix} \underline{f}_C \\ \underline{f}_I \end{pmatrix} \quad , \tag{10}$$

geschrieben werden kann. Die Matrix K_I besitzt die Blockdiagonalgestalt

$$K_I = diag\,(K_{I,i})_{i=1,2,\ldots,p} \quad , \tag{11}$$

wobei die Diagonalblöcke $K_{I,i}$ die Dimension $N_{I,i} \times N_{I,i}$ haben und die betrachtete partielle Differentialgleichung in Ω_i unter reinen homogenen Dirichletschen Randbedingungen auf $\partial\Omega_i$ approximieren, d.h.

$$(K_{I,i}\underline{u}_{I,i}, \underline{v}_{I,i}) = a(\Phi V_{I,i}\underline{u}_{I,i}, \Phi V_{I,i}\underline{v}_{I,i}) \quad \forall \underline{u}_{I,i}, \underline{v}_{I,i} \in \mathbf{R}^{N_{I,i}} \tag{12}$$

mit $V_{I,i}^T = (O \; I_{I,i} \; O)_{N_{I,i} \times N}$ und $\mathbf{V}_{I,i} = span(\Phi V_{I,i}) \subset \overset{\circ}{\mathbf{W}}{}_2^1(\Omega_i)$ bzw. $\mathbf{V}_{I,i} \subset [\overset{\circ}{\mathbf{W}}{}_2^1(\Omega_i)]^n$ im Falle eines Differentialgleichungssystems aus n partiellen Differentialgleichungen. Da die Basisfunktionen ϕ_i für alle $i = N_C + 1, \cdots, N$ auf Γ_C verschwinden, kann

$$\mathbf{V}_I = \mathbf{V}_{I,1} \oplus \cdots \oplus \mathbf{V}_{I,p} \tag{13}$$

als orthogonale Summe der Unterräume $V_{I,1}, \cdots, V_{I,p}$ bzgl. des energetischen Skalarproduktes $a(\cdot, \cdot)$ geschrieben werden. Durch die Identität (8) werden die anfangs erwähnten Eigenschaften der Bilinearform $a(\cdot, \cdot)$ auf die Steifigkeitsmatrix K übertragen. Insbesondere ist K symmetrisch und positiv definit.

Wir führen nun neben der Knotenbasis Φ die sogenannte diskrete harmonische Basis

$$\overset{*}{\Psi} = \left\{ \overset{*}{\psi}_1, \cdots, \overset{*}{\psi}_{N_C}, \overset{*}{\psi}_{N_C+1}, \cdots, \overset{*}{\psi}_N \right\} = \Phi \overset{*}{V} \tag{14}$$

von $\mathbf{V} = span(\Phi) = span(\overset{*}{\Psi})$ ein. Die Basistransformationsmatrix $\overset{*}{V} = \begin{pmatrix} \overset{*}{V}_C & \overset{*}{V}_I \end{pmatrix}$ wird durch die Teilmatrizen

$$\overset{*}{V}_C = \begin{pmatrix} I_C \\ -K_I^{-1} K_{IC} \end{pmatrix}_{N \times N_C} \quad \text{und} \quad \overset{*}{V}_I = \begin{pmatrix} O \\ I_I \end{pmatrix}_{N \times N_I} \tag{15}$$

eindeutig beschrieben. Daraus ist ersichtlich, daß alle Basisfunktionen, die zu den inneren Knoten der Teilgebiete $\Omega_1, \ldots, \Omega_p$ gehören, unverändert bleiben, d.h. $\overset{*}{\psi}_j = \phi_j$ für alle $j = N_C + 1, \ldots, N$. Analog zu (4) kann der F.E. Teilraum $\mathbf{V}$ in der Form

$$\mathbf{V} = \overset{*}{\mathbf{V}}_C \oplus \overset{*}{\mathbf{V}}_I \tag{16}$$

dargestellt werden, wobei $\overset{*}{\mathbf{V}}_C = span(\Phi \overset{*}{V}_C)$ und $\overset{*}{\mathbf{V}}_I = span(\Phi \overset{*}{V}_I) = \mathbf{V}_I$. Die Zerlegung (16) ist offenbar orthogonal bezüglich des energetischen Skalarproduktes $a(\cdot,\cdot)$. In der neuen Basis $\overset{*}{\Psi}$ nimmt das F.E. Gleichungssystem die Gestalt

$$\begin{pmatrix} S_C & O \\ O & K_I \end{pmatrix} \begin{pmatrix} \overset{*}{\underline{u}}_C \\ \overset{*}{\underline{u}}_I \end{pmatrix} = \begin{pmatrix} \underline{f}_C - K_{CI}K_I^{-1}\underline{f}_I \\ \underline{f}_I \end{pmatrix} \tag{17}$$

an. Der erste Diagonalblock

$$S_C = K_C - K_{CI}K_I^{-1}K_{IC} \tag{18}$$

wird als Schur-Komplement bezeichnet. Der Lösungsvektor $\overset{*}{\underline{u}} = \left(\overset{*}{\underline{u}}_C^T, \overset{*}{\underline{u}}_I^T\right)^T$ ist nun nichts anderes als der Koeffizientenvektor bezüglich der diskreten harmonischen Basis $\overset{*}{\Psi}$, d.h. $u = \overset{*}{\Psi}\, \overset{*}{\underline{u}} = \Phi\, \overset{*}{V}\, \overset{*}{\underline{u}} \in \mathbf{V}$ ist genau die Lösung des F.E. Schemas (6). Den Lösungsvektor $\underline{u} = \left(\underline{u}_C^T, \underline{u}_I^T\right)^T$ in der Knotenbasis Φ erhalten wir durch die Rücktransformation

$$\underline{u} = \overset{*}{V}\, \overset{*}{\underline{u}} = \begin{pmatrix} I_C & O \\ -K_I^{-1}K_{IC} & I_I \end{pmatrix} \begin{pmatrix} \overset{*}{\underline{u}}_C \\ \overset{*}{\underline{u}}_I \end{pmatrix} = \begin{pmatrix} \overset{*}{\underline{u}}_C \\ \overset{*}{\underline{u}}_I - K_I^{-1}K_{IC}\, \overset{*}{\underline{u}}_C \end{pmatrix} \tag{19}$$

Die Beziehungen (17) - (19) sind offensichtlich äquivalent zu den wohlbekannten Formeln

$$\left.\begin{array}{rcl} \overset{*}{\underline{u}}_I &=& K_I^{-1}\underline{f}_I \,, \\ \underline{g}_C &=& \underline{f}_C - K_{CI}\, \overset{*}{\underline{u}}_I \,, \quad S_C \;=\; K_C - K_{CI}K_I^{-1}K_{IC} \,, \\ \underline{u}_C &=& \overset{*}{\underline{u}}_C \;=\; S_C^{-1}\underline{g}_C \,, \\ \underline{u}_I &=& \overset{*}{\underline{u}}_I - K_I^{-1}K_{IC}\, \overset{*}{\underline{u}}_C \end{array}\right\} \tag{20}$$

der klassischen F.E. Substrukturtechnik. Wegen der Blockdiagonalgestalt (11) von K_I ist die Realisierung von K_I^{-1} gleichbedeutend mit der parallelen Lösung von p voneinander unabhängigen kleineren Gleichungssystemen mit den Systemmatrizen $K_{I,1}, \ldots, K_{I,p}$. Obwohl die Lösung dieser Gleichungssysteme ohne jegliche Kommunikation vollständig parallel durchgeführt werden kann, ist die direkte Lösung eines einzelnen Teilsystems in der Praxis oft zu aufwendig. Deshalb wenden wir nicht die klassische F.E. Substrukturtechnik direkt an, sondern lösen das F.E. Gleichungssystem (10) mit einer durch Gebietszerlegung parallelisierten und vorkonditionierten Version des konjugierten Gradientenverfahrens. Diese parallelisierte Version stellen wir im nächsten Abschnitt vor.

3 Eine parallelisierte Version des CG-Verfahrens

Die in diesem Abschnitt beschriebene Parallelisierungsidee geht auf K.H. Law [Law89] zurück und wurde später für verschiedene Vorkonditionierungstechniken von den Autoren weiterentwickelt [HL91a, HLM90, HLM91, Mey90]. Die auf der Gebietszerlegungsmethode beruhende Parallelisierung des CG-Verfahrens eignet sich besonders für eine Implementierung auf Mehrprozessorrechnern mit lokalem Speicher und Botschaftenaustausch über Links.

Entsprechend den p Teilgebieten $\overline{\Omega}_i \ \ (i = 1, 2, \ldots, p)$ verteilen wir alle Matrizen und Vektoren, die im CG-Algorithmus vorkommen, auf die p Prozessoren $P_i \ \ (i = 1, 2, \ldots, p)$ unseres Mehrprozessorrechners. Für die Vektoren nutzen wir zwei Typen der Verteilung, die wir überlappend (Typ I) und addierend (Typ II) nennen wollen:

Typ I : $\underline{\mathbf{u}}, \underline{\mathbf{w}}$ und $\underline{\mathbf{s}}$ werden im Prozessor P_i ($\hat{=} \overline{\Omega}_i$) als $\underline{\mathbf{u}}_i = A_i \underline{\mathbf{u}}$, $\underline{\mathbf{w}}_i = A_i \underline{\mathbf{w}}$ und $\underline{\mathbf{s}}_i = A_i \underline{\mathbf{s}}$ gespeichert;

Typ II : $\underline{r}, \underline{v}, \underline{f}$ werden in P_i als $\underline{r}_i, \underline{v}_i, \underline{f}_i$ gespeichert, so daß $\underline{r} = \sum\limits_{i=1}^{p} A_i^T \underline{r}_i$ etc. gilt,

wobei A_i die Superelementzusammenhangsmatrix von $\overline{\Omega}_i$ ist, d.h. die Matrix A_i ist eine Boolesche Matrix der Dimension $N_i \times N$ und bildet einen Vektor $\underline{g} \in \mathbf{R}^N$, der alle Knotenparameter enthält, in einen Vektor $\underline{g}_i \in \mathbf{R}^{N_i}$ ab, der nur die zu $\overline{\Omega}_i$ gehörenden Knotenparameter beinhaltet. Mit den eingeführten Booleschen Matrizen A_i kann nun die Steifigkeitsmatrix K in der Form

$$K = \sum_{i=1}^{p} A_i^T K_i A_i \qquad (21)$$

dargestellt werden, wobei K_i die zum Superelement $\overline{\Omega}_i$ gehörende Superelementsteifigkeitsmatrix ist. In Analogie dazu benutzen wir die Notationen

$$K_{IC,i} = (K_i)_{IC} \ , \ A_{C,i} = (A_i)_C \quad \text{etc.}$$

Mit der oben eingeführten Datenverteilung und den entsprechenden Bezeichnungen können wir nun die parallelisierte Version des CG-Algorithmus wie folgt aufschreiben.

For all $i = 1, 2, \ldots, p$ do in parallel

0. Startschritt

Bestimme eine Anfangsnäherung $\underline{\mathbf{u}}_i$, z.B., $\underline{\mathbf{u}}_i = O$

$$\begin{array}{lcl}
\underline{r}_i & = & \underline{f}_i - K_i \underline{\mathbf{u}}_i \\
\underline{\mathbf{w}} & = & C^{-1} \sum A_i^T \underline{r}_i \\
\underline{\mathbf{s}}_i & = & \underline{\mathbf{w}}_i \qquad (\underline{\mathbf{w}}_i = A_i \underline{\mathbf{w}}) \\
\sigma_i & = & \underline{\mathbf{w}}_i^T \underline{r}_i \\
\sigma & = & \sigma^0 = \sum \sigma_i
\end{array}$$

Iteration

$$\begin{array}{llcl}
1. & \underline{\check{v}}_i & = & K_i \underline{\check{\mathbf{s}}}_i \\
 & \delta_i & = & \underline{\check{v}}_i^T \underline{\check{\mathbf{s}}}_i \\
 & \delta & = & \sum \delta_i \\
 & \alpha & = & \check{\sigma} \, / \, \delta \\
2. & \underline{\mathbf{u}}_i & = & \underline{\check{\mathbf{u}}}_i + \alpha \underline{\check{\mathbf{s}}}_i \\
 & \underline{r}_i & = & \underline{\check{r}}_i - \alpha \underline{\check{v}}_i \\
3. & \underline{\mathbf{w}} & = & C^{-1} \sum A_i^T \underline{r} \\
4. & \sigma_i & = & \underline{\mathbf{w}}_i^T \underline{r}_i \\
 & \sigma & = & \sum \sigma_i \\
 & \beta & = & \sigma / \check{\sigma} \\
5. & \underline{\mathbf{s}}_i & = & \underline{\mathbf{w}}_i + \beta \underline{\check{\mathbf{s}}}_i \\
6. & \sigma & \leq & \varepsilon^2 * \sigma^0 \quad ? \ \xrightarrow{\text{nein}} \ \text{goto 1.}
\end{array}$$

$\downarrow$ ja

stop

Im oben aufgeschriebenen Algorithmus markiert das Symbol " ˇ " die vorangegangenen Iterierten. Das Summenzeichen $\sum = \sum\limits_{i=1}^{p}$ bedeutet eine Aufsummierung bzw. Typenumwandlung, die immer mit Transferoperationen (Kommunikation !) verbunden ist.

Selbst im Falle $C = I$, dh. ohne Vorkonditionierung, erfordert der Schritt 3

$$\underline{\mathbf{w}} = C^{-1} \sum A_i^T \underline{r} \qquad (22)$$

bzw. der entsprechende Startschritt eine Umwandlung eines Vektors vom Typ II in einen Vektor vom Typ I und ist folglich mit einem lokalen Datenaustausch vom Umfang $O(h^{-(d-1)})$ Daten über benachbarte Teilgebietskoppelränder hinweg verbunden. Von einem Vorkonditionierungsoperator C werden wir deshalb unter anderem verlangen, daß der Kommunikationsaufwand gegenüber dem nichtvorkonditionierten Fall $C = I$ nicht erhöht wird. Vorkonditionierungsoperatoren, die dieser Forderung entsprechen, stellen wir im Abschnitt 4 dieser Arbeit vor.

Im Schritt 6 des oben aufgeschriebenen Algorithmus wird der Genauigkeitstest in der $KC^{-1}K$-energetischen Norm $\|\cdot\|_{KC^{-1}K} = (KC^{-1}K\cdot,\cdot)^{0.5}$ durchgeführt. Für gute Vorkonditionierungsoperatoren C, wie wir sie im Auge haben, ist dieser Test dem in der numerischen Praxis noch oft anzutreffenden Defektnormtest (= Test in der K^2-energetischen Norm) unbedingt vorzuziehen (vgl. dazu [GL90]). In der Theorie des CG-Verfahrens wird der relative Fehler jedoch in der K-energetischen Norm abgeschätzt (siehe z.B. [SN89]). Wenn die Spektraläquivalenzungleichungen

$$\underline{\gamma} C \;\leq\; K \;\leq\; \overline{\gamma} C \tag{23}$$

mit gewissen positiven Spektraläquivalenzkonstanten $\underline{\gamma}$ und $\overline{\gamma}$ gelten, dann erhält man die Fehlerabschätzung

$$\| \underline{\mathbf{u}} - \underline{\mathbf{u}}^k \|_K \leq\; q_k \;\| \underline{\mathbf{u}} - \underline{\mathbf{u}}^0 \|_K \tag{24}$$

wobei $\underline{\mathbf{u}} = \left(\underline{\mathbf{u}}_C^T, \underline{\mathbf{u}}_I^T\right)^T$ die exakte Lösung des F.E. Gleichungssystems (7) bzw. (10) ist, $\underline{\mathbf{u}}^k$ bezeichnet die k-te Iterierte des CG-Algorithmus, $\|\cdot\|_K = (\cdot,\cdot)_K^{0.5} = (K\cdot,\cdot)^{0.5}$ ist die K-energetische Norm, $\xi = \underline{\gamma} \,/\, \overline{\gamma}$, $\rho = (1-\xi^{0.5}) \,/\, (1+\xi^{0.5})$, und $q_k = 2\rho^k \,/\, (1+\rho^{2k})$. Folglich wird der Anfangsfehler $\underline{\mathbf{u}} - \underline{\mathbf{u}}^0$ durch $k = \sharp \ln(\varepsilon^{-1} + (\varepsilon^{-2}+1)^{0.5}) \,/\, \ln(\rho^{-1}) \sharp$ Iterationen mindestens um den Faktor ε $(0 < \varepsilon < 1)$ reduziert. Hierbei bezeichnet $\sharp x \sharp$ die kleinste ganze Zahl, die größer oder gleich x ist. Die Ungleichungen (23) werden im Sinne des Euklidischen Skalarproduktes verstanden.

4 Vorkonditionierung durch Gebietszerlegung (DD)

4.1 Die einfache DD-Vorkonditionierung

Die im Abschnitt 2 beschriebene Transformation der Knotenbasis Φ in die diskret harmonische Basis $\overset{*}{\Psi}$ mit der Basistransformationsmatrix $\overset{*}{V}$ hat die Beziehung

$$diag\,(S_C, K_I) = \overset{*}{V}{}^T K \,\overset{*}{V} \tag{25}$$

zwischen der Systemmatrix $diag\,(S_C, K_I)$ in der Basis $\overset{*}{\Psi}$ und der Systemmatrix K in der Basis Φ zur Folge. Aus (25) erhalten wir unmittelbar die Faktorisierung

$$\begin{aligned} K &= \overset{*}{V}{}^{-T}\, diag\,(S_C, K_I)\, \overset{*}{V}{}^{-1} \\ &= \begin{pmatrix} I_C & K_{CI}K_I^{-1} \\ O & I_I \end{pmatrix} \begin{pmatrix} S_C & O \\ O & K_I \end{pmatrix} \begin{pmatrix} I_C & O \\ K_I^{-1}K_{IC} & I_I \end{pmatrix} \end{aligned} \tag{26}$$

der Steifigkeitsmatrix K.
Ersetzt man nun formal S_C und K_I durch geeignete, symmetrische und positiv definite Vorkonditionierungsoperatoren C_C und $C_I = diag\,(C_{I,i})_{i=1,\dots,p}$, dann erhält man die sogenannte einfache DD-Vorkonditionierung

$$C = \begin{pmatrix} I_C & K_{CI}C_I^{-1} \\ O & I_I \end{pmatrix} \begin{pmatrix} C_C & O \\ O & C_I \end{pmatrix} \begin{pmatrix} I_C & O \\ C_I^{-1}K_{IC} & I_I \end{pmatrix}, \tag{27}$$

die zunächst von C. Börgers [B8̈9b] und später von A. Meyer [Mey90] und in einer gemeinsamen Arbeit von G. Haase und U. Langer [HL91b] untersucht wurde. Von den Autoren wurde in [HLM91] der folgende Satz bewiesen

Satz 4.1 *Falls die symmetrischen und positiv definiten Vorkonditionierungsoperatoren C_C und C_I den Spektraläquivalenzungleichungen*

$$\underline{\gamma}_C C_C \le S_C \le \overline{\gamma}_C C_C \quad \text{und} \quad \underline{\gamma}_I C_I \le K_I \le \overline{\gamma}_I C_I \tag{28}$$

mit gewissen positiven Konstanten $\underline{\gamma}_C, \overline{\gamma}_C, \underline{\gamma}_I$ und $\overline{\gamma}_I$ genügen, dann gelten für die einfache DD-Vorkonditionierung (27) die Spektraläquivalenzungleichungen (23) mit den Spektraläquivalenzkonstanten

$$\begin{aligned} \underline{\gamma} &= \underline{\gamma}_1 \left(1 + \tfrac{1}{2}\left(\mu - (\mu^2 + 4\mu)^{0.5}\right)\right) \qquad \text{und} \\ \overline{\gamma} &= \overline{\gamma}_1 \left(1 + \tfrac{1}{2}\left(\mu + (\mu^2 + 4\mu)^{0.5}\right)\right) \end{aligned}$$

wobei $\underline{\gamma}_1 = \min\{\underline{\gamma}_C, \underline{\gamma}_I\}$, $\overline{\gamma}_1 = \max\{\overline{\gamma}_C, \overline{\gamma}_I\}$ *und* $\mu = \rho(S_C^{-1} T_C)$ *ist, mit* $T_C = K_{CI}(K_I^{-1} - C_I^{-1})K_I(K_I^{-1} - C_I^{-1})K_{IC}$, *gleich dem Spektralradius von* $S_C^{-1}T_C$. *Darüber hinaus gilt für die relative spektrale Konditionszahl* $\kappa(C^{-1}K)$ *die zweiseitige Abschätzung*

$$\begin{aligned} &\left(\underline{\gamma}_1 \,/\, \overline{\gamma}_1\right)\left(1 + \frac{1}{2}\left(\mu + (\mu^2 + 4\mu)^{0.5}\right)\right)^2 \\ &\quad \le \kappa(C^{-1}K) \le \left(\overline{\gamma}_1 \,/\, \underline{\gamma}_1\right)\left(1 + \tfrac{1}{2}\left(\mu + (\mu^2 + 4\mu)^{0.5}\right)\right)^2 . \end{aligned} \tag{29}$$

Die zweiseitige Abschätzung (29) sagt aus, daß diese Abschätzung in einem gewissen Sinne scharf ist [HLM91]. Falls $\overline{\gamma}_1 \,/\, \underline{\gamma}_1 = O(1)$ für $h \longrightarrow 0$, dann verhält sich $\kappa(C^{-1}K)$ wie μ^2.

Die einfache DD-Vorkonditionierung (27) erfüllt auch unsere Forderung aus Abschnitt 3 bzgl. der Kommunikation. Tatsächlich die Vorkonditionierungsgleichung in Schritt 3 bzw. im entsprechenden Startschritt des parallelisierten CG-Algorithmus kann man nun in der Form

$$\left.\begin{aligned} \underline{\mathbf{w}}_C &= C_C^{-1} \sum A_{C,i}^T \left(\underline{r}_{C,i} - K_{CI,i} C_{I,i}^{-1} \underline{r}_{I,i}\right) \\ \underline{\mathbf{w}}_{I,i} &= C_{I,i}^{-1} \left(\underline{r}_{I,i} - K_{IC,i} \underline{\mathbf{w}}_{C,i}\right) \end{aligned}\right\} \tag{30}$$

schreiben. Der kritische Punkt in (30) ist allerdings die Vorkonditionierung C_C für S_C. Bei der Wahl von C_C ist darauf zu achten, daß kein oder nur ein geringer Kommunikationsaufwand hinzukommt (siehe Abschnitt 5).

O. Axelsson und P.S. Vassilevski nutzen ebenfalls die Faktorisierung $K^{-1} = \overset{*}{V} \, diag\left(S_C^{-1}, K_I^{-1}\right) \overset{*}{V}^T$ zur Konstruktion von iterativen Lösungsverfahren für das F.E. Gleichungssystem (10), [AV91].

4.2 Die ASM-DD-Vorkonditionierung

Wir ersetzen die Matrix K_I in der Basistransformationsmatrix $\overset{*}{V}$ durch eine geeignet gewählte nichtsinguläre Matrix $B_I = diag\,(B_{I,i})_{i=1,\dots,p}$, von der wir zunächst voraussetzen, daß sie die gleiche Blockdiagonalstruktur wie K_I hat, aber nichtnotwendig symmetrisch und positiv definit ist. Gleichungssysteme mit den Matrizen $B_{I,i}$ sollen jedoch wesentlich effektiver lösbar sein als Systeme mit den ursprünglichen Matrizen $K_{I,i}$ (siehe z.B. [HLM90] und Abschnitt 5). Die damit erhalte Basis

$$\widetilde{\Psi} = \left\{\widetilde{\psi}_1, \cdots, \widetilde{\psi}_{N_C}, \widetilde{\psi}_{N_C+1}, \cdots, \widetilde{\psi}_N\right\} = \Phi \widetilde{V} \tag{31}$$

bezeichnen wir in Anlehnung an die Basis $\overset{*}{\Psi}$ als näherungsweise diskrete harmonische Basis. Die Basistransformationsmatrix $\widetilde{V} = \left(\widetilde{V}_C \; \widetilde{V}_I\right)$ wird jetzt durch die Teilmatrizen

$$\widetilde{V}_C = \begin{pmatrix} I_C \\ -B_I^{-1} K_{IC} \end{pmatrix}_{N \times N_C} \qquad \text{and} \quad \widetilde{V}_I = \overset{*}{V}_I = V_I = \begin{pmatrix} O \\ I_I \end{pmatrix}_{N \times N_I} \quad . \tag{32}$$

bestimmt. Die zu den inneren Knoten der Teilgebiete gehörenden Basisfunktionen bleiben wiederum unverändert, d.h. $\widetilde{\psi}_j = \varphi_j$ für alle $j = N_C + 1, \ldots, N$. Der F.E. Raum

$$\mathbf{V} \;=\; \widetilde{\mathbf{V}}_C \;+\; \widetilde{\mathbf{V}}_I \tag{33}$$

Kann nun als direkte Summe der Teilräume $\widetilde{\mathbf{V}}_I = span(\Phi \widetilde{V}_I)$ und $\widetilde{\mathbf{V}}_C = span(\Phi \widetilde{V}_C)$ dargestellt werden. Im Vergleich zur diskreten harmonischen Basis verlieren wir zwar die Orthogonalität bezüglich des energetischen Skalarproduktes $a(\cdot,\cdot)$, aber bei geeigneter Wahl von B_I kann erwartet werden, daß der Winkel $\angle(\widetilde{V}_C, \widetilde{V}_I)$ zwischen den Räumen $\widetilde{V}_C$ und $\widetilde{V}_I$ nahe der 90°-Marke ist. Tatsächlich, in [HLM90] haben wir das folgende Lemma bewiesen.

Lemma 4.1 *Der Winkel zwischen den Räumen $\widetilde{V}_C$ und $\widetilde{V}_I$ ist durch die Beziehung*

$$\gamma = \cos \angle(\widetilde{V}_C, \widetilde{V}_I) = \sqrt{\frac{\mu}{1+\mu}} \; , \tag{34}$$

gegeben, wobei $S_C = K_C - K_{CI} K_I^{-1} K_{IC}$ das Schur-Komplement, T_C der Operator $T_C = K_{CI}(K_I^{-1} - B_I^{-T}) K_I (K_I^{-1} - B_I^{-1}) K_{IC}$ und $\mu = \rho(S_C^{-1} T_C)$ der Spektralradius von $S_C^{-1} T_C$ sind.

Folglich kann für die Unterräume $\widetilde{\mathbf{V}}_C$ und $\widetilde{\mathbf{V}}_I$ die verschärfte Cauchy-Ungleichung

$$|a(u,v)| \leq \gamma \; |||\, u \,||| \; |||\, v \,||| \quad \forall u \in \widetilde{\mathbf{V}}_C \; , \; \forall v \in \widetilde{\mathbf{V}}_I \tag{35}$$

mit der nicht verbesserbaren Konstanten $\gamma \in [\,0,1)$ aus (34) aufgeschrieben werden. Hier und im weiteren wird mit $|||\cdot||| = (a(\cdot,\cdot))^{0.5}$ die Energienorm bezeichnet.

Mit der Darstellung (33) ist die Additive Schwarzsche Methode (ASM)

$$u^{k+1} = u^k + P_C z^k + P_I z^k \quad ; \quad k = 0, 1, \ldots \; , \tag{36}$$

verknüpft, wobei $u^0 \in \mathbf{V}$ eine beliebige Startnäherung und $z^k = u - u^k$ den k-ten Iterationsfehler bezeichnen [BMS91, DW87, DW90, DW91, HLM90, Xu90]. Die Orthoprojektoren $P_I \, : \, \mathbf{V} \longrightarrow \widetilde{\mathbf{V}}_I$ und $P_C \, : \, \mathbf{V} \longrightarrow \widetilde{\mathbf{V}}_C$ werden entsprechend durch die Beziehungen

$$a(P_C u, v) = a(u,v) \quad \forall u \in \mathbf{V} \; , \forall v \in \widetilde{\mathbf{V}}_C \tag{37}$$

und

$$a(P_I u, v) = a(u,v) \quad \forall u \in \mathbf{V} \; , \forall v \in \widetilde{\mathbf{V}}_I \tag{38}$$

definiert. Die ASM (36) in $\mathbf{V} = \widetilde{\mathbf{V}}_C + \widetilde{\mathbf{V}}_I$ ist äquivalent zum Verfahren der einfachen Iteration

$$\tau^{-1} D \, (\underline{u}^{k+1} - \underline{u}^k) + K \, \underline{u}^k = \underline{f} \quad ; \quad k = 0, 1, \ldots \tag{39}$$

in $\mathbf{R}^N$ bzgl. der Knotenbasis Φ, wobei $u^k = \Phi \underline{u}^k$, $\tau = 1$, $D = \widetilde{V}^{-T} \widetilde{D} \widetilde{V}^{-1}$ und $\widetilde{D} = diag(S_C + T_C, K_I)$. Aus Lemma 4.1 erhalten wir unmittelbar die Spektraläquivalenzungleichungen

$$(1-\gamma)\, D \;\leq\; K \;\leq\; (1+\gamma)\, D \, , \tag{40}$$

mit den nicht verbesserbaren Spektraläquivalenzkonstanten $1-\gamma$ und $1+\gamma$. Folglich ist die relative spektrale Konditionszahl $\kappa(D^{-1}K)$ durch die Beziehung

$$\kappa(D^{-1}K)=\frac{\lambda_{max}(D^{-1}K)}{\lambda_{min}(D^{-1}K)}=\frac{1+\gamma}{1-\gamma}=\left(\sqrt{\mu}+\sqrt{1+\mu}\right)^2 \tag{41}$$

gegeben. Aus (40) folgt rückwirkend, daß der optimale Iterationsparameter $\tau_{opt}=\ 2\,/\,(1+\gamma+1-\gamma)=1$ tatsächlich gleich 1 und die optimale Rate $q_{opt}=(1+\gamma-(1-\gamma))\,/\,\ (1+\gamma+1-\gamma)$ gleich γ sind. Wir ersetzen nun die Diagonalblöcke S_C+T_C und K_I in $\widetilde{D}$ durch geeignet gewählte, symmetrische und positiv definite Vorkonditionierungsoperatoren C_C und C_I und erhalten somit die sogenannte ASM-DD-Vorkonditionierung

$$C=\begin{pmatrix} I_C & K_{CI}B_I^{-T}\\ O & I_I\end{pmatrix}\begin{pmatrix} C_C & O\\ O & C_I\end{pmatrix}\begin{pmatrix} I_C & O\\ B_I^{-1}K_{IC} & I_I\end{pmatrix}\ , \tag{42}$$

die in [HLM90] ausführlich untersucht wurde. In eben dieser Arbeit wurde der folgende Satz bewiesen.

Satz 4.2 *Falls die symmetrischen und positiv definiten Vorkonditionierungsoperatoren C_C und C_I den Spektraläquivalenzungleichungen*

$$\underline{\gamma}_C\, C_C\ \le\ S_C+T_C\ \le\ \overline{\gamma}_C\, C_C \quad \textit{und} \quad \underline{\gamma}_I\, C_I\ \le\ K_I\ \le\ \overline{\gamma}_I\, C_I \tag{43}$$

mit gewissen positiven Konstanten $\underline{\gamma}_C,\overline{\gamma}_C,\underline{\gamma}_I$ und $\overline{\gamma}_I$ genügen, dann gelten für die einfache ASM-DD-Vorkonditionierung (42) die Spektraläquivalenzungleichungen (23) mit den Spektraläquivalenzkonstanten

$$\begin{aligned}\underline{\gamma}&=\min\{\underline{\gamma}_C,\underline{\gamma}_I\}\left(1-\sqrt{\tfrac{\mu}{1+\mu}}\right) \qquad \textit{und}\\ \overline{\gamma}&=\max\{\overline{\gamma}_C,\overline{\gamma}_I\}\left(1+\sqrt{\tfrac{\mu}{1+\mu}}\right)\end{aligned} \tag{44}$$

wobei $\mu=\rho(S_C^{-1}T_C)$, mit $T_C=K_{CI}(K_I^{-1}-B_I^{-T})K_I(K_I^{-1}-B_I^{-1})K_{IC}$, gleich dem Spektralradius von $S_C^{-1}T_C$ ist. Darüber hinaus gilt für $\kappa(C^{-1}K)$ die zweiseitige Abschätzung

$$\begin{aligned}&\left(\underline{\gamma}_1\,/\,\overline{\gamma}_1\right)\left(\sqrt{\mu}+\sqrt{1+\mu}\right)^2\\ &\quad\le\ \kappa(C^{-1}K)\ \le\ \left(\overline{\gamma}_1\,/\,\underline{\gamma}_1\right)\left(\sqrt{\mu}+\sqrt{1+\mu}\right)^2\end{aligned} \tag{45}$$

mit $\underline{\gamma}_1=\min\{\underline{\gamma}_C,\underline{\gamma}_I\}$ und $\overline{\gamma}_1=\max\{\overline{\gamma}_C,\overline{\gamma}_I\}$.

Falls $\overline{\gamma}_1\,/\,\underline{\gamma}_1=O(1)$ für $h\longrightarrow 0$, dann wird das Verhalten von $\kappa(C^{-1}K)=O(\mu)$ ausschließlich durch das Verhalten des Spektralradius $\mu=\rho(S_C^{-1}T_C)$ für $h\longrightarrow 0$ bestimmt. Die Zahl μ kann ihrerseits nur durch die Wahl der Basistransformationskomponente B_I beeinflußt werden. Im Gegensatz zur Vorkonditionierung aus Abschnitt 4.1 ist jetzt B_I von C_I abgekoppelt.

Die Vorkonditionierungsgleichung in Schritt 3 bzw. im entsprechenden Startschritt des parallelisierten CG-Algorithmus aus Abschnitt 3 kann man für die ASM-DD-Vorkonditionierung (42) nun in der Form

$$\left.\begin{aligned}\check{\underline{\mathbf{w}}}_C\ &=\ C_C^{-1}\,\sum\,A_{C,i}^T\left(\check{\underline{\mathrm{r}}}_{C,i}-K_{CI,i}B_{I,i}^{-T}\check{\underline{\mathrm{r}}}_{I,i}\right)\\ \check{\underline{\mathbf{w}}}_{I,i}\ &=\ C_{I,i}^{-1}\check{\underline{\mathrm{r}}}_{I,i}-B_{I,i}^{-1}K_{IC,i}\check{\underline{\mathbf{w}}}_{C,i}\end{aligned}\right\} \tag{46}$$

schreiben.

4.3 Die MSM-DD-Vorkonditionierung

Unter Benutzung der in (37) und (38) definierten Orthoprojektoren P_C und P_I können wir die klassische Multiplikative Schwarzsche Methode (MSM) in der Form

$$u^1 = u^0 + P_I\, z^0 \quad , \quad u^{k+1} = (I - P)\, u^k + Pu \quad , \quad k = 1, 2, \dots \,, \tag{47}$$

schreiben, wobei $P = P_I + P_C - P_I P_C$, $z^0 = u - u^0$ und $u^0 \in \mathbf{V}$ ein beliebig gewähltes Startelement ist [BMS91, DW87, DW90, DW91, HLM90, HL91a, Xu90]. Für den Fehler $z^k = u - u^k$ erhalten wir aus (36) sofort die Beziehung

$$z^1 = (I - P_I)\, z^0 \quad , \quad z^{k+1} = (I - P_I)\,(I - P_C)\, z^k \quad , \quad k = 1, 2, \dots \,. \tag{48}$$

Die Orthoprojektionen in die Unterräume $\widetilde{\mathbf{V}}_C$ und $\widetilde{\mathbf{V}}_I$ sind algebraisch äquivalent zur Lösung von Gleichungssystemen mit den Sytemmatrizen $S_C + T_C$ und K_I. In der Praxis ist die exakte Lösung dieser Gleichungssysteme i.allg. zu aufwendig. Deshalb verwenden wir zur Lösung dieser Gleichungssysteme schnellkonvergente Iterationsverfahren und starten diese mit dem Nullvektor als Anfangsnäherung. Aus der symmetrischen Version

$$z^{k+1} = (I - P_I)\,(I - P_C)\,(I - P_I)\, z^k \quad , \quad k = 0, 1, \dots \tag{49}$$

der MSM werden in [HL91a] die Ergebnisse des folgenden Satzes abgeleitet.

Satz 4.3 *Der in der Basistransformationsmatrix $\widetilde{V}$ enthaltene Operator B_I sei durch die Beziehung*

$$B_I = K_I\, \left(I_I - \overline{M_I}^r\right)^{-1} \tag{50}$$

und die Blockvorkonditionierungsoperatoren C_I und C_C durch die Beziehungen

$$C_I = K_I\, (I_I - M_I^s)^{-1} \quad \text{und} \quad C_C = (S_C + T_C)\, \left(I_C - Q_C^t\right)^{-1} \tag{51}$$

definiert, d.h. sowohl B_I als auch C_I und C_C werden implizit durch r, s und t Iterationsschritte stationärer Iterationsverfahren mit den Iterationsoperatoren $\overline{M_I}^r$, M_I^s und Q_C^t bestimmt. Die Iterationsoperatoren $\overline{M_I}^r$, M_I^s und Q_C^t mögen die folgenden Voraussetzungen erfüllen:

1. *M_I and Q_C sind selbstadjungiert im K_I- bzw. $(S_C{+}T_C)$- energetischen Skalarprodukt,*

2. $$\begin{aligned} \mu &= \rho(S_C^{-1} T_C) = \sup_{\underline{v} \in \overset{*}{V}_C \mathrm{R}^{N_C} \setminus \{O\}} \left\| \begin{pmatrix} O & O \\ O & \overline{M_I}^r \end{pmatrix} \underline{v} \right\|_K^2 \Big/ \| \, \underline{v} \, \|_K^2 \\ &= \left\| \begin{pmatrix} O & O \\ O & \overline{M_I} \end{pmatrix}^r \Bigg|_{\overset{*}{V}_C \mathrm{R}^{N_C}} \right\|_K^2 \le \overline{\mu}\,, \end{aligned} \tag{52}$$

3. $$\| \, M_I \, \|_{K_I} = \varrho(M_I) \le \eta < 1. \tag{53}$$

4. $$\| \, Q_C \, \|_{S_C + T_C} = \varrho(Q_C) \le \nu < 1\,. \tag{54}$$

Dann ist die symmetrische MSM (49) mit den durch (51) definierten näherungsweisen Orthoprojektoren äquivalent zur Methode der einfachen Iteration

$$\underline{u}^{k+1} = (I - C^{-1} K)\, \underline{u} + C^{-1} \underline{f} \quad , \quad k = 0, 1, \dots \tag{55}$$

mit der MSM-DD-Vorkonditionierung

$$
\begin{aligned}
C \;=\; & \begin{pmatrix} I_C & K_{CI}(I_I - \overset{*}{\overline{M_I}}{}^r M_I^s)K_I^{-1} \\ O & I_I \end{pmatrix} * \begin{pmatrix} C_C & O \\ O & K_I(I_I - M_I^{2s})^{-1} \end{pmatrix} \\
& * \begin{pmatrix} I_C & O \\ (I_I - M_I^s \overline{M_I}{}^r)K_I^{-1}K_{IC} & I_I \end{pmatrix} ,
\end{aligned} \tag{56}
$$

wobei $\overset{*}{\overline{M_I}}$ *der bzgl. des* K_I*-energetischen Skalarproduktes zu* $\overline{M_I}$ *adjungierte Iterationsoperator ist, dh.*

$$
(\overline{M_I}\underline{u}_I, \underline{v}_I)_{K_I} = (\underline{u}_I, \overset{*}{\overline{M_I}}\underline{v}_I)_{K_I} \quad \forall \underline{u}_I, \underline{v}_I \in \mathbf{R}^{N_I} . \tag{57}
$$

mit $(\cdot,\cdot)_{K_I} = (K_I\cdot,\cdot)$. *Die Matrix* C *ist symmetrisch und positiv definit. Desweiteren gelten die Spektraläquivalenzungleichungen*

$$
(1-q)\,C \;\leq\; K \;\leq\; (1+q)\,C , \tag{58}
$$

mit

$$
q = \nu^t + (1-\nu^t)\left(\eta^s + (1-\eta^s)\sqrt{\frac{\overline{\mu}}{(1+\overline{\mu})}}\right)^2 \;<\; 1. \tag{59}
$$

Bemerkung 4.1 *Für ebene elliptische RWA zweiter Ordnung gilt die grobe Abschätzung*

$$
\overline{\mu} \;\leq\; \overline{\eta}^{2r}(\overline{c}_2\underline{c}_1^{-1}h^{-1} - 1)
$$

wobei $\underline{c}_1$ *und* $\overline{c}_2$ *h-unabhängige, positive Konstanten aus den wohlbekannten Ungleichungen*

$$
\underline{c}_1 h I_C \leq S_C \leq \overline{c}_1 I_C \qquad \text{und} \qquad \underline{c}_2 I_C \leq K_C \leq \overline{c}_2 I_C
$$

sind und $\overline{\eta} < 1$ *eine obere Schranke für* $\| \overline{M_I} \|_{K_I}$ *ist [B\"89b, HLM90, HLM91, HL91b]. Eine h-unabhängige obere Schranke wird aus der obigen Abschätzung für* $\overline{\mu}$ *nur dann erreicht, wenn* $r = O(\ln h^{-1})$ *und die Rate* $\overline{\eta}$*, wie typischerweise im Falle von Multigrid-Verfahren, h-unabhängig ist. In der Praxis sind aber ein oder zwei Multigrid-Iterationen (d.h.* $r = 1$ *oder 2) hinreichend, um de facto ein h-unabhängiges* $\overline{\mu}$ *zu erreichen (vgl. numerische Resultate im Abschnitt 5.1 und in [HLM90, HLM91, HL91b, JL] sowie Schlußfolgerung 5 in Kapitel 5 aus [HLM91]).*

Damit ist die symmetrische MSM (49) mit den durch (51) definierten näherungsweisen Orthoprojektoren nichts anderes als die ASM (vgl. (39) und (42))

$$
\tau^{-1}C\,(\underline{u}^{k+1} - \underline{u}^k) + K\underline{u}^k = \underline{f} \quad , \quad k = 0,1,\ldots \tag{60}
$$

mit $\tau = 1$ und der ASM-DD-Vorkonditionierung (42), deren Komponenten B_I, C_I und C_C nun durch die Beziehungen

$$
\left.
\begin{array}{lcll}
B_I & = & K_I\,(I_I - M_I^s\overline{M_I}{}^r)^{-1} & , \\
C_I & = & K_I\,(I_I - M_I^{2s})^{-1} & \text{und} \\
C_C & = & (S_C + T_C)\,(I_C - Q_C^t)^{-1} &
\end{array}
\right\} \tag{61}
$$

definiert werden. Von dieser quivalenz kann man zweifach profitieren. Zum einen gelten die Spektraläquivalenzungleichungen (58) - (59), die auf der MSM-Ratenabschätzung beruhen [HL91a]. Zum anderen kann die spezielle Form (56) der MSM-DD-Vorkonditionierung C bei der Lösung des Vorkonditionierungssystems $C\,\underline{w} = \underline{r}$ im CG-Verfahren ausgenutzt werden. Tatsächlich, die Vorkonditionierungsgleichung im Schritt 3 bzw. im entsprechenden Startschritt des parallelisierten CG-Algorithmus aus Abschnitt 3 kann man für die MSM-DD-Vorkonditionierung (56) in der Form

$$
\begin{array}{llcl}
1) & \underline{w}^s_{I,i} & = & (I_{I,i} - M^s_{I,i})\, K^{-1}_{I,i} \underline{r}_{I,i} \\
2) & \underline{w}^{s+r}_{I,i} & = & B^{-T}_{I,i} \underline{r}_{I,i} \; = \overset{*}{\overline{M_{I,i}}}{}^{r} \underline{w}^s_{I,i} + (I_{I,i} - \overset{*}{\overline{M_{I,i}}}{}^{r})\, K^{-1}_{I,i} \underline{r}_{I,i} \\
3) & \underline{w}_C & = & C_C^{-1} \sum A^T_{C,i} (\underline{r}_{C,i} - K_{CI,i} \underline{w}^{s+r}_{I,i}) \\
4) & \underline{w}^r_{I,i} & = & (I_{I,i} - \overline{M_{I,i}}{}^{r})\, K^{-1}_{I,i} (-K_{IC,i} \underline{w}_{C,i}) \\
5) & \underline{\tilde{r}}_{I,i} & = & \underline{r}_{I,i} - K_{IC,i} \underline{w}_{C,i} \\
6) & \underline{\tilde{w}}_{I,i} & = & \underline{w}^s_{I,i} + \underline{w}^r_{I,i} \\
7) & \underline{w}_{I,i} & = & M^s_{I,i} \underline{\tilde{w}}_{I,i} + (I_{I,i} - M^s_{I,i})\, K^{-1}_{I,i} \underline{\tilde{r}}_{I,i}
\end{array}
$$

geschrieben werden [HL91a]. Die Gleichung 1) ist gleichbedeutend mit der Anwendung von s $M_{I,i}$-Iterationsschritten (d.h. s Iterationsschritten des Iterationsverfahrens, das durch den Iterationsoperator $M_{I,i}$ definiert wird) auf das Gleichungssystem

$$K_{I,i}\, \underline{w}_{I,i} \;=\; \underline{r}_{I,i} \tag{62}$$

mit dem Nullvektor $\underline{w}^0_{I,i} = O$ als Startvektor, während 2) die Anwendung von r $\overset{*}{\overline{M_{I,i}}}$-Iterationschritten auf die gleiche Gleichung (62) aber mit der Startnäherung $\underline{w}^s_{I,i}$ bedeutet. In Analogie dazu sind die Beziehungen 4) und 7) jeweils äquivalent zur Anwendung von r $\overline{M_{I,i}}$-Iterationschritten auf die Gleichung

$$K_{I,i}\, \underline{w}_{I,i} \;=\; -K_{IC,i} \underline{w}_{C,i}$$

mit der Startnäherung $\underline{w}^0_{I,i} = O$ bzw. von s $M_{I,i}$-Iterationsschritten auf die Gleichung

$$K_{I,i}\, \underline{\tilde{w}}_{I,i} \;=\; \underline{\tilde{r}}_{I,i}$$

mit der Startlösung $\underline{\tilde{w}}_{I,i}$. Folglich sparen wir $2s$ Iterationsschritte des durch den selbstadjungierten Iterationsoperator $M_{I,i}$ definierten Iterationsverfahrens ein.

4.4 Die hierarchische DD-Vorkonditionierung

Im Unterschied zu den Abschnitten 4.1 bis 4.3 ist die Gebietszerlegung für die in diesem Abschnitt zu beschreibende, sehr effektive und einfach realisierbare hierarchische Vorkonditionierung nur Mittel zum Zwecke der Parallelisierung. Analog zu den Abschnitten 4.1 bis 4.3 läßt sich aus der Gebietszerlegungsidee eine besonders elegante Parallelisierung ableiten, die die von uns geforderte Eigenschaft (einmaliger Austausch von Koppelrandinformationen pro Iterationsschritt) beibehält. Diese parallele Realisierung auf der Basis der Gebietszerlegung ist auch Grund dafür, daß wir den Namen hierarchische DD-Vorkonditionierung gewählt haben, obwohl sie algebraisch vollständig äquivalent zur hierarchischen Vorkonditionierung von H. Yserentant ist, die inzwischen Eingang in die numerische Praxis gefunden hat [Yse86b, Yse86a, Yse90]. Die theoretischen Fakten werden deshalb hier nur kurz skizziert und können mit den obigen Veröffentlichungen detailliert nachvollzogen werden.

Wie oft bei Multigrid-ähnlichem Herangehen wird von einer Geometriebeschreibung des Gebietes Ω in Form einer Grobvernetzung (Nutzer-Triangularisierung) ausgegangen (vgl. auch Bild 1). Diese (möglichst wenigen) groben Dreiecke sollen die Geometrie von Ω einschließlich deren Interfacestruktur ausreichend genau beschreiben (eventuell durch krummlinige Dreiecke). Außerdem wird vorausgesetzt, daß die Teildreiecke zur Definition der Teilgebiete Ω_i verwendet werden. Somit wäre eine Anzahl von Grobdreiecken besonders günstig, die mit einem Vielfachen der Prozessoranzahl übereinstimmt. Wir nennen die Knoten dieser Startvernetzung Knoten vom Level 0 und ordnen jedem solchen Knoten eine übliche F.E. Basisfunktion $\hat{\psi}_i = \phi_{h_0,i}$ zu, die als Knotenbasis zur Grobvernetzung zu verstehen ist. Im weiteren wird eine hierarchische Verfeinerung des Grobnetzes in etwa l Verfeinerungsstufen vorgenommen. Im einfachsten Falle derart, daß ein Dreieck des Levels $j-1$ durch Verbinden der Seitenmitten in 4 ähnliche Dreiecke des Levels j geteilt wird (weitere Verfeinerungstechniken siehe [JL]). Den im Level j neu hinzugekommenen Knoten werden wieder die üblichen Knotenbasisfunktionen $\psi_i = \phi_{h_j,i}$ des Levels j zur bisherigen Basis hinzugefügt, so daß am Ende dieses Prozesses N_0 Knotenbasisfunktionen vom Level 0, N_1 weitere vom Level 1 usw. definiert wurden und folglich die Dimension $N = N_0 + N_1 + \ldots + N_l$ ist. Die Gesamtheit dieser Basisfunktionen nennt man hierarchische Basis

$$\hat{\Psi} = \left\{ \hat{\psi}_1, \cdots, \hat{\psi}_{N_0}, \hat{\psi}_{N_0+1}, \cdots, \hat{\psi}_N \right\} \tag{63}$$

des F.E. Raumes $\mathbf{V} = span(\hat{\Psi}) = span(\Phi)$. Würde man anstelle der üblichen Knotenbasis Φ der feinsten Vernetzung mit dieser hierarchischen Basis $\hat{\Psi}$ arbeiten, dann würden sich folgende zwei interessante Fakten ergeben [Mey90]:

1. Die zugehörige Steifigkeitsmatrix $\widehat{K} = \left(a(\hat{\Psi}_j, \hat{\Psi}_i) \right)_{i,j=1,\ldots,N}$ ist fast voll besetzt.

2. Die Konditionszahl von $\widehat{K}$ ist relativ klein, $\kappa(\widehat{K}) = O(l^2) = O(|\ln h|^2)$.

Wegen der zweiten Eigenschaft wäre das Verfahren der konjugierten Gradienten (ohne Vorkonditionierung !) ein sehr schneller Löser für ein Gleichungssystem mit der Systemmatrix $\widehat{K}$, da nur $O(l)$ Iterationsschritte für eine bestimmte fixierte relative Genauigkeit notwendig würden. Dem steht die erste Eigenschaft gegenüber, die eine Speicherung von $\widehat{K}$ und eine schnelle Multiplikation der Art $\widehat{K} * \underline{v}$ zunächst als nicht opportun erscheinen läßt. Wieder entsteht aus der Basistransformation

$$\hat{\Psi} = \Phi \hat{V} \tag{64}$$

mit einer noch zu untersuchenden, nichtsingulären $(N \times N)$-Basistransformationsmatrix $\hat{V}$ der elegante Ausweg aus dieser Situation. Aus (64) folgt

$$\widehat{K} = \hat{V}^T K \hat{V} \tag{65}$$

und somit

$$\kappa(\widehat{K}) = \kappa(\hat{V}^T K \hat{V}) = \kappa(\hat{V} \hat{V}^T K) \ ,$$

also wäre die Vorkonditionierungsmatrix

$$C^{-1} = \hat{V} \hat{V}^T$$

eine sehr effektive Variante im vorkonditionierten CG-Verfahren für das übliche F.E.-Gleichungssystem $K \, \underline{u} = \underline{f}$ in der Knotenbasis Φ. Von besonderer Wichtigkeit ist hierbei die Tatsache, daß die beiden Matrixmultiplikationen

$$\underline{w} = \hat{V} \hat{V}^T \underline{r}$$

im Vorkonditionierungsschritt des CG-Verfahrens (siehe Abschnitt 3) außerordentlich effektiv realisierbar sind (jeweils N Multiplikationen und $2N$ Additionen). Hierzu ist lediglich eine sogenannte hierarchische Liste aller Knotenzusammenhänge notwendig, in der geordnet vom niedrigsten zum höchsten Level alle Knotennummern des jeweiligen Levels verzeichnet sind.

Bedenkt man nun die Parallelisierung des gesamten Rechenprozesses durch Gebietszerlegung (DD), bedeutet dies, daß jeder Prozessor P_i nur einige Dreiecke des Grobnetzes erhält und seinen Teil des hierarchischen Netzes für $\overline{\Omega_i}$ generiert. Damit ist auch auf jedem Prozessor nur der lokale Anteil der hierarchischen Liste über die Knotenentstehung bekannt. Trotzdem ist dies der hierarchischen Vorkonditionierung nicht hinderlich, da die beiden Multiplikationsalgorithmen

$$\underline{y} = \widehat{V}^T \underline{r} \qquad \text{und} \qquad \underline{w} = \widehat{V} \underline{y}$$

im Zusammenhang mit der verteilten Speicherung von $\underline{r}$, $\underline{w}$ und der hierarchischen Liste für $\widehat{V}$ folgende Eigenschaften haben. Das Residium $\underline{r}$ ist nach Typ II (siehe Abschnitt 3)

$$\underline{r} = \sum_{i=1}^{p} A_i^T \underline{r}_i$$

über die Prozessoren verteilt. Die lokale Multiplikation

$$\underline{y}_i = \widehat{V}_i^T \underline{r}_i$$

mit den nur lokal bekannten Knotenzusammenhängen erhält die Typ II -Eigenschaft für $\underline{y}$, d.h. es gilt wieder

$$\underline{y} = \sum_{i=1}^{p} A_i^T \underline{y}_i \, .$$

Wird dieser Akkumulationsschritt jetzt ausgeführt und ein Vektor $\underline{\widetilde{y}}$ $(= \underline{y})$ vom Typ I erzeugt, d.h.

$$\underline{\widetilde{y}}_j = \underline{y}_j + \sum_{\substack{i=1 \\ i \neq j}}^{p} A_i^T \underline{y}_i \, ,$$

so ändert die zweite lokale Multiplikation

$$\underline{w}_i = \widehat{V}_i \, \underline{\widetilde{y}}_i$$

abermals nichts an der Typ I -Speichertechnik für $\underline{w}$. Somit benötigt ein CG-Schritt mit hierarchischer Vorkonditionierung genau den selben Kommunikationsaufwand wie <u>ohne</u> jede Vorkonditionierung (vgl. Abschnitt 3 mit $C = I$) und lediglich $4N_i$ wesentliche Operationen zusätzlich.

Bemerkung 4.2 *Zum Ausgleich von starken Schwankungen in den Koeffizienten der Dgl. ist eine Jacobi-Modifikation der Art*

$$C^{-1} = J^{-\frac{1}{2}} \widehat{V} \, \widehat{V}^T J^{-\frac{1}{2}}$$

mit $J = diag\, K$ *bzw.* $J = blockdiag\, K$ *zu empfehlen.*

Bemerkung 4.3 *Die von H. Yserentant hervorgehobene zusätzliche Grobgitterlösung zwischen der $\widehat{V}$- und $\widehat{V}^T$-Multiplikation bringt für die Parallelisierung zusätzliche Probleme der Organisation und des Transfers und ist deshalb von uns weggelassen worden.*

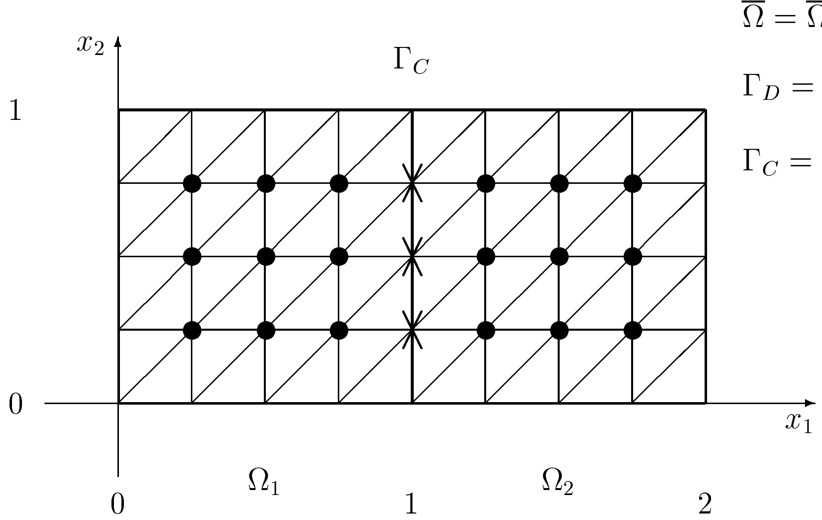
Bemerkung 4.4 *Für 3D-Probleme kann die Vorgehensweise ungeändert übernommen werden. Natürlich ist die Konstruktion hierarchischer 3D-Vernetzungen wesentlich komplizierter (vgl. etwa [B89a]). Der Hauptnachteil besteht allerdings in der wesentlichen Verschlechterung der Konditionszahl von $\widehat{K}$ (bzw. von $C^{-1}K = \widehat{V}\widehat{V}^T K$) zu $O(h^{-1})$. Durch eine Abänderung des Vorkonditionierers unter Verwendung von L_2-Projektoren kann dies zwar behoben werden [BPX90, Yse90], allerdings wird sich eventuell der Kommunikationsbedarf drastisch erhöhen.*

5 Numerische Experimente

5.1 Numerische Experimente mit ASM- und MSM-DD-Vorkonditionierungen

In den Arbeiten [HLM90, HLM91, HL91a, JL] wurden eine Vielzahl von numerischen Experimenten mit der ASM-DD-Vorkonditionierung dargestellt und ausgewertet. Sowohl zur Basistransformation B_I als auch zur K_I-Vorkonditionierung C_I haben wir dabei Multigrid-Techniken verwendet, d.h. die Operatoren B_I und C_I wurden durch die Beziehungen (50) und (51) definiert, wobwi jetzt $\overline{M_I}$ und M_I entsprechende Multigrid-Iterationsoperatoren sind. Zur Bestimmung der Blockvorkonditionierung C_C wurden in Abhängigkeit vom Aussehen des Koppelrandes Γ_C (ohne oder mit Kreuzungspunkten) die Vorschläge von M. Dryja [Dry82] (ohne Kreuzungspunkte) und J.M. Bramble, J.E. Pasciak, A.H. Schatz [BPS89] (mit Kreuzungspunkten) in einer leicht modifizierten Form sowie eine in [HLM90] vorgeschlagene hierarchische Technik aufgegriffen. Als Modellbeispiel diente uns das Dirichlet-Problem für die Poisson-Gleichung in verschiedenen Gebieten.

In diesem Abschnitt wollen wir einige vergleichende numerische Ergebnisse zur Nutzung der ASM- und MSM-DD-Vorkonditionierung darstellen. Als Modellbeispiel (= Beispiel 1 in den Arbeiten [HLM90, HLM91, HL91a, JL]) benutzten wir der Einfachheit halber wiederum das Dirichlet-Problem für die Poisson-Gleichung im Rechteck $\overline{\Omega} = [\,0,2\,] \times [\,0,1\,] = \overline{\Omega}_1 \cup \overline{\Omega}_2$, das in die zwei Einheitsquadrate $\overline{\Omega}_1 = (\,0,1\,) \times (\,0,1\,)$ und $\overline{\Omega}_1 = (\,1,2\,) \times (\,0,1\,)$ zerlegt wurde (siehe Bild 2).



$\overline{\Omega} = \overline{\Omega}_1 \cup \overline{\Omega}_2$

$\Gamma_D = \partial\Omega$

$\Gamma_C = \{\,(1,y) : 0 < y < 1\,\}$

Abbildung 2: Gebietszerlegung und Triangularisierung (2 Gitter) des Modellgebietes Ω ($N_C = 3,\ N_I = 18,\ N_{I,1} = N_{I,2} = 9$)

Die skalierte Dryjasche Vorkonditionierung

$$C_C \;=\; \delta \begin{pmatrix} 2 & -1 & & & O \\ -1 & 2 & -1 & & \\ & \ddots & \ddots & \ddots & \\ & & -1 & 2 & -1 \\ O & & & -1 & 2 \end{pmatrix}^{0.5}_{N_C \times N_C}$$

mit $\delta = 2$ diente uns in allen Fällen als $(S_C + T_C)$-Blockvorkonditionierung. Die Basistransformation B_I und K_I-Vorkonditionierung wurden wie in den vorangegangenen Arbeiten [HLM90, HLM91, HL91a] durch Multigrid-Techniken definiert. In den Tabellen 1-7 bedeutet z.B. $B_I = V11$ (ω-Jacobi), daß der V-Zyklus mit 1 ω-Jacobi-Schritt ($\omega = 2/3$ in allen Fällen) als Vorglättung und 1 ω-Jacobi-Schritt als Nachglättung benutzt wird; $C_I = W22$ (Gauß-Seidel $\uparrow\uparrow\downarrow\downarrow$) bedeutet, daß der W-Zyklus mit 2 lexikographisch vorwärtigen Gauß-Seidel-Schritten als Vorglättung und 2 lexikographisch rückwärtigen Gauß-Seidel-Schritten als Nachglättung verwendet wurde usw.. In allen Beispielrechnungen wurde die lineare Interpolation und die kanonisch dazugehörige Restriktion (= transponiert zur Interpolation) benutzt. Sowohl in B_I als auch in C_I wurde nur ein Multigrid-Schritt durchgeführt, d.h. $r = s = 1$ (siehe auch (50) und (51) sowie Bemerkung (4.1)). Die Boxen der Tabellen 1-7 enthalten folgende Informationen

	Gitter ↓	
ASM- /MSM- DD Vorkonditionierung ⟶	Anzahl der Iterationen	$x * 10^{-7}$ (Relativer Fehler) / Durchschnittliche CG-Rate

wobei der relative Fehler in der $KC^{-1}K$-energetischen Norm gemessen wurde. Die Abbruchschranke war in allen Beispielen $\varepsilon = 10^{-6}$.

Anzahl der Gitter	3		4		5		6	
ASM	10	9.84 / 0.251	13	2.20 / 0.308	15	4.87 / 0.379	21	6.51 / 0.507
MSM	7	6.88 / 0.132	8	2.40 / 0.149	9	2.94 / 0.188	11	3.53 / 0.259

Tabelle 1: $C_I = V11(\omega\text{-Jacobi})$, $B_I = V11(\omega\text{-Jacobi})$, $C_C = \delta(-1\ 2\ -1)^{0.5}$

Anzahl der Gitter	3		4		5		6	
ASM	9	2.52 / 0.185	11	5.80 / 0.271	14	7.43 / 0.365	20	7.40 / 0.494
MSM	6	3.75 / 0.085	6	3.96 / 0.086	6	5.77 / 0.091	7	1.38 / 0.105

Tabelle 2: $C_I = V11$ (Gauß-Seidel $\uparrow\downarrow$), $B_I = V11(\omega\text{-Jacobi})$, $C_C = \delta(-1\ 2\ -1)^{0.5}$

Anzahl der Gitter	3		4		5		6	
ASM	9	9.85 / 0.215	10	6.51 / 0.241	11	3.65 / 0.260	12	4.70 / 0.297
MSM	7	3.49 / 0.120	7	4.87 / 0.125	7	5.85 / 0.129	7	9.17 / 0.137

Tabelle 3: $C_I = V11(\omega\text{-Jacobi})$, $B_I = V11$ (Gauß-Seidel $\uparrow\downarrow$) , $C_C = \delta(-1\ 2\ -1)^{0.5}$

Bereits diese einfachen Beispiele zeigen eine deutliche Verbesserung der Konvergenzeigenschaften des CG-Verfahrens durch die Benutzung der MSM-DD-Vorkonditionierung im Vergleich zur ASM-DD-Vorkonditionierung.

Anzahl der Gitter	3		4		5		6	
ASM	8	$\frac{2.31}{0.148}$	8	$\frac{5.09}{0.163}$	9	$\frac{3.78}{0.193}$	10	$\frac{4.54}{0.232}$
MSM	6	$\frac{2.70}{0.080}$	6	$\frac{1.51}{0.073}$	5	$\frac{9.91}{0.063}$	6	$\frac{1.21}{0.070}$

Tabelle 4: $C_I = V11$ (Gauß-Seidel $\uparrow\downarrow$), $B_I = V11$ (Gauß-Seidel $\uparrow\downarrow$) , $C_C = \delta(-1\ 2\ -1)^{0.5}$

Anzahl der Gitter	3		4		5		6	
ASM	7	$\frac{8.33}{0.135}$	9	$\frac{3.67}{0.193}$	12	$\frac{5.53}{0.201}$	17	$\frac{4.57}{0.424}$
MSM	6	$\frac{1.12}{0.069}$	5	$\frac{3.30}{0.051}$	5	$\frac{1.78}{0.045}$	5	$\frac{0.97}{0.040}$

Tabelle 5: $C_I = W22$ (Gauß-Seidel $\uparrow\uparrow\downarrow\downarrow$) , $B_I = V11(\omega$-Jacobi), $C_C = \delta(-1\ 2\ -1)^{0.5}$

Anzahl der Gitter	3		4		5		6	
ASM	9	$\frac{3.34}{0.191}$	9	$\frac{7.91}{0.210}$	9	$\frac{6.29}{0.205}$	9	$\frac{6.81}{0.206}$
MSM	7	$\frac{2.01}{0.110}$	7	$\frac{2.28}{0.112}$	7	$\frac{1.85}{0.109}$	7	$\frac{9.10}{0.098}$

Tabelle 6: $C_I = V11(\omega$-Jacobi), $B_I = W22$ (Gauß-Seidel $\uparrow\uparrow\downarrow\downarrow$) , $C_C = \delta(-1\ 2\ -1)^{0.5}$

Anzahl der Gitter	3		4		5		6	
ASM	6	$\frac{6.74}{0.094}$	6	$\frac{3.40}{0.084}$	6	$\frac{1.90}{0.076}$	6	$\frac{1.97}{0.076}$
MSM	5	$\frac{5.16}{0.055}$	5	$\frac{2.17}{0.046}$	4	$\frac{9.05}{0.031}$	4	$\frac{3.45}{0.024}$

Tabelle 7: $C_I = W22$ (Gauß-Seidel $\uparrow\uparrow\downarrow\downarrow$) , $B_I = W22$ (Gauß-Seidel $\uparrow\uparrow\downarrow\downarrow$) $C_C = \delta(-1\ 2\ -1)^{0.5}$

5.2 Numerische Experimente mit der hierarchischen DD-Vorkonditionierung auf einem Transputerhypercube

Aus den umfangreichen numerischen Experimenten mit der hierarchischen DD-Vorkonditionierungstechnik auf Parallelrechnern, die aus 2 bis 32 Transputern T800 (Hypercubearchitektur) aufgebaut waren, sollen hier zwei typische Beispiele ausgewählt werden.

Zunächst betrachten wir die Poisson-Gleichung

$$-\Delta u = -1$$

mit wechselnden Dirichlet- und Neumann-Randbedingungen in dem ringförmigen Gebiet Ω aus Abschnitt 2 (siehe Bild 1). Die Grobtriangularisierung besteht aus 16 Dreiecken, so daß 2,4,8 und 16 Prozessoren zum Einsatz kamen. Bei gegebener Gebietszerlegung hat die Zuordnung der Dreiecke zu den Prozessoren keinen Einfluß auf die Güte der hierarchischen Vorkonditionierung (im Gegensatz zur DD-Vorkonditionierung). Deshalb ergeben sich bei allen Prozessoranzahlen gleiche Iterationszahlen, die in der Tabelle 8 angegeben werden.

Levels l	Anzahl der Unbekannten N	Anzahl der Iterationen ($\varepsilon = 10^{-4}$)	Rechenzeiten [in Sekunden] $p = 2$	$p = 4$	$p = 8$	$p = 16$
2	160	23	3”	3”	3”	3”
3	576	28	4”	3”	3”	3”
4	2175	34	8”	5”	5”	3”
5	8448	39	28”	16”	10”	7”
6	33280	44	-	60”	33”	20”
7	132096	49	-	-	125”	70”

Tabelle 8: Numerische Resultate für $-\Delta u = -1$ im ringförmigen Gebiet Ω (Bild 1) unter wechselnden Dirichlet- und Neumann-Randbedingungen (siehe auch Bemerkung 5.1)

Bemerkung 5.1 *Die ersten Zeilen der Tabelle 8 zeigen, daß von allen Zeiten etwa 3” zu subtrahieren sind (Zwischenergebnisprotokollierung). Hiermit zeigen die unteren Zeilen in der Tat Beschleunigungen nahezu proportional zur Prozessoranzahl.*

Zu Vergleichszwecken wählen wir nun ein etwas komplizierteres Problem. Wir betrachten ein ebenes (ebener Verzerrungszustand), lineares Elastizitätsproblem im gleichen ringförmigen Gebiet Ω aus Bild 1 mit der gleichen Anordnung der Dirichlet- und Neumann-Ränder. Obwohl im Vergleich zur Poisson-Gleichung jetzt doppelt so viele Unbekannte zu berechnen sind und die Matrix-Vektor-Multiplikation aufwendiger ist, erhöhen sich die Rechenzeiten bei der hierarchischen DD-Vorkonditionierung (jetzt für jede Verschiebungskomponente einzeln angewendet !) nur um das 3-fache, wie folgende Beispielrechnung mit 8 Prozessoren zeigt (Tabelle 9).

Im Bild 3 ist das deformierte Gebiet zu sehen, was im Ergebnis einer Rechnung auf dem Level 3 erhalten wurde.

Um die Prozessoranzahl im Hypercube innerhalb der Hypercube-Architektur weiter erhöhen zu können, wurden die Transputer über ein Link zu Doppelknoten verbunden. Diese Doppelknoten mit jetzt 6 freien Links nach außen können bis zu einem 6-dimensionalen Hypercube vernetzt werden. Für eine Gebietszerlegung in 32 Grobdreiecke wird deshalb eine Beispielrechnung für ein ebenes, lineares Elastizitätsproblem mit 8,16 und 32 Prozessoren (vernetzt als 2-,3- und 4-dimensionaler Hypercube von Doppelknoten) angefügt. In diesem Beispiel ist $\Omega = (0,1) \times (0,1)$ das Einheitsquadrat, das in 4 mal 4

Levels l	Anzahl der Unbekannten N	Anzahl der Iterationen ($\varepsilon = 10^{-4}$)	Rechenzeiten [in Sekunden] $p = 8$
2	320	35	3"
3	1152	40	5"
4	4352	44	10"
5	16896	46	30"
6	66560	49	113"

Tabelle 9: Numerische Resultate für ein lineares Elastizitätsproblem im ringförmigen Gebiet Ω (Bild 2) unter wechselnden Verschiebungs- und Kräfterandbedingungen

Quadrate und somit in 32 kongruente Dreiecke zerlegt wurde. Vorgegebene Verschiebungen waren $\vec{u} = (0,0)^T$ bei $y = 0$ und $\vec{u} = (0, -\delta)^T$ bei $y = 1$ und $0.25 \leq x \leq 0.75$. Die Tabelle 10 zeigt wieder Iterationszahlen und Rechenzeiten für dieses Beispiel. Abbbruchschranke ε war in allen Beispielen 10^{-4}. Das deformierte Gebiet ist im Bild 4 dargestellt.

Level l	Anzahl der Unbekannten N	Anzahl der Iterationen ($\varepsilon = 10^{-4}$)	Rechenzeiten [in Sekunden] $p = 8$	$p = 16$	$p = 32$
3	2178	40	6"	4"	4"
4	8450	42	13"	8"	7"
5	33282	45	41"	23"	17"
6	132096	47	-	84"	53"
7	526338	48	-	-	185"

Tabelle 10: Numerische Resultate für ein lineares Elastizitätsproblem im Einheitsquadrat unter wechselnden Randbedingungen

Literatur

[AV91] Ove Axelsson and P. S. Vassilevski. Construction of variable-step preconditioners for inner-outer iteration methods. In *Proceedings of the IMACS*, 1991. Brussels, April 2-4, 1991.

[B8̈9a] E. Bänsch. Local mesh refinement in 2 and 3 dimensions. Technical Report SFB256, Rhein. Friedrich-Wilhelm-Univ. Bonn, 1989.

[B8̈9b] C. Börgers. The Neumann-Dirichlet domain decomposition method with inexact solvers on the subdomains. *Numer. Math*, 55(2):123–136, 1989.

[BMS91] B. E. Bjørstad, R. Moe, and M. Skogen. Parallel domain decomposition and iterative refinement algorithms. In *[Hac91]*, pages 28–46, 1991.

[BPS89] J. H. Bramble, J. E. Pasciak, and A. H. Schatz. The construction of preconditioners for elliptic problems by substructuring. i-iv. *Math. Comp.*, 47, 49, 51, 53(175, 179):103–234, 1–16, 415–430, 1–24, 1986, 1987, 1988, 1989.

Abbildung 3: Das zum undeformierten Gebiet Ω aus Bild 1 gehörende deformierte Gebiet nach einer Berechnung des Elastizitätsproblems auf dem Level 3

Abbildung 4: Deformiertes Einheitsquadrat beim Elastizitätsproblem (siehe Tabelle 10).

[BPX90] J. H. Bramble, J. E. Pasciak, and J. Xu. Parallel multilevel preconditioners. *Math. Comp.*, 1990.

[CGPW89] Tony F. Chan, R. Glowinski, J. Périaux, and O. B. Widlund, editors. *Second International Symposium on Domain Decomposition Methods for Partial Differential Equations*, Philadelphia, 1989. SIAM. Los Angeles, California, January 14-16, 1988.

[CGPW90] Tony F. Chan, R. Glowinski, J. Périaux, and O. B. Widlund, editors. *Third International Symposium on Domain Decomposition Methods for Partial Differential Equations*, Philadelphia, 1990. SIAM. Houston, March 20-22, 1989.

[Cia78] Ph. Ciarlet. *The finite element method for elliptic problems.* North-Holland Publishing Company, Amsterdam-New York-Oxford, 1978.

[Dry82] M. Dryja. A capacitance matrix method for Dirichlet problems on polygonal regions. *Numer. Math.*, 39(1):51–64, 1982.

[DW87] M. Dryja and O. B. Widlund. An additive variant of Schwarz alternating method for the case of many subregions. Technical Report 339, Department of Computer Science, Courant Institute, New York, 1987. also Ultracomputer Note 131.

[DW90] M. Dryja and O. B. Widlund. Towards a unified theory of domain decomposition algorithms for elliptic problems. In *[CGPW90]*, pages 3–21. SIAM, Philadelphia, 1990.

[DW91] M. Dryja and O. B. Widlund. Multilevel additive methods for elliptic finite element problems. In *[Hac91]*, pages 58–69, 1991.

[GGMP88] R. Glowinski, G. H. Golub, G. A. Meurant, and J. Périaux, editors. *First International Symposium on Domain Decomposition Methods for Partial Differential Equations*, Philadelphia, 1988. SIAM. Paris, January 1987.

[GL90] Gerhard Globisch and Ulrich Langer. On the use of multigrid preconditioners in a multigrid software package. In Sabine Hengst, editor, *Proceedings of the "4-th Multigrid Seminar"held at Unterwirbach, GDR, May 2-6, 1989*, pages 105–134, Berlin, 1990. Academy of Science. Report-Nr. R-MATH-03/90.

[Hac91] W. Hackbusch, editor. *Parallel Algorithms for Partial Differential Equations*, Braunschweig, 1991. Vieweg. Proceedings of the Sixth GAMM-Seminar, Kiel, January 19-21, 1990.

[HL91a] Gundolf Haase and Ulrich Langer. The non-overlapping domain decomposition multiplicative Schwarz method. Preprint, TU Chemnitz, 1991. Eingereicht.

[HL91b] Gundolf Haase and Ulrich Langer. On the use of multigrid preconditioners in the domain decomposition method. In *[Hac91]*, pages 101–110, 1991.

[HLM90] Gundolf Haase, Ulrich Langer, and Arnd Meyer. A new approach to the Dirichlet decomposition method. In Sabine Hengst, editor, *Proceedings of the "5-th Multigrid Seminar"held at Eberswalde, GDR, May 14-18, 1990*, pages 1–59, Berlin, 1990. Academy of Science. Report-Nr. R-MATH-09/90.

[HLM91] Gundolf Haase, Ulrich Langer, and Arnd Meyer. Domain decomposition preconditioners with inexact subdomain solvers. Preprint 192, TU Chemnitz, 1991. Appears in the "Journal of Numerical Linear Algebra with Applications".

[JL] Michael Jung and Ulrich Langer. Applications of multilevel methods to practical problems. *Surveys on Mathematics for Industry.* eingereicht.

[Law89] K. H. Law. A parallel finite element solution method. *Computer and Structures*, 23(6):845–858, 1989.

[Lio90] P. L. Lions. On the Schwarz alternating method I,II,III. In *[GGMP88], [CGPW89], [CGPW90]*, pages 1–42, 47–70, 202–223. SIAM, Philadelphia, 1988, 1989, 1990.

[Mey90] Arnd Meyer. A parallel preconditioned conjugate gradient method using domain decomposition and inexact solvers on each subdomain. *Computing*, 45:217–234, 1990.

[Prz63] J. S. Przemieniecki. Matrix structural analysis of substructures. *AIAA J.*, 1:138–147, 1963.

[Sch90] H. A. Schwarz. *Über einige Abbildungsaufgaben. Gesammelte Mathemathische Abhandlungen*, volume 2. Springer, Bonn, 1890. First published in "Vierteljahresschrift der Naturforschenden Gesellschaft in Zürich"1870, v.15, pp 272-286.

[SN89] A. A. Samarski and E. S. Nikolaev. *Numerical methods for grid equations. Vol. II : Iterative Methods.* Birkhäuser-Verlag, Basel-Boston-Berlin, 1989.

[Sob36] S. L. Sobolev. The Schwarz algorithm in the theory of elasticity. *Dokl. Acad. Nauk USSR*, 4:243–246, 1936.

[Xu90] J. Xu. Iterative methods by space decomposition and subspace correction: A unifying approach. Technical Report AM67, Department of Mathematics, Penn State University, 1990.

[Yse86a] H. Yserentant. Hierarchical basis give conjugate gradient type methods a multigrid speed of convergence. *Appl. Math. and Comp.*, 19:347–358, 1986.

[Yse86b] H. Yserentant. On the multi-level splitting of finite element spaces. *Numer. Math.*, 49(4):379–412, 1986.

[Yse90] H. Yserentant. Two preconditioners on the multi-level-splitting of finite element spaces. *Numer. Math.*, 58:163–184, 1990.