Electronic Transactions on Numerical Analysis.
Volume 15, pp. 132-151, 2003.

ISSN 1068-9613.

ETNA
Kent State University
etna@mcs.kent.edu

A MULTIGRID ALGORITHM FOR SOLVING THE MULTI-GROUP, ANISOTROPIC SCATTERING BOLTZMANN EQUATION USING FIRST-ORDER SYSTEM LEAST-SQUARES METHODOLOGY*

B. CHANG AND B. LEE †

Abstract. This paper describes a multilevel algorithm for solving the multi-group, anisotropic scattering Boltzmann equation formulated with a first-order system least-squares methodology. A $P_n - h$ finite element discretization is used. The resulting angle discretization of this P_n approach does not exhibit the so-called "ray effects," but this discretization leads to a large coupled system of partial differential equations for the spatial coefficients, and, on scaling the system to achieve better approximation, the system coupling depends strongly on the material parameters. Away from the thick, low absorptive regime, a relatively robust multigrid algorithm for solving these spatial systems will be described. For the thick, low absorptive regime, where an incompressible elasticity-like equation appears, an additive/multiplicative Schwarz smoother gives substantial multigrid improvement over standard nodal smoothers. Rather than using higher-order or Raviart-Thomas finite element spaces, which lead to complicated implementation, only low-order, conforming finite elements are used. Numerical examples illustrating almost $h-$independent convergence rates and locking-free discretization accuracy will be given.

Key words. Boltzmann transport equation, first-order system least-squares, multigrid method.

AMS subject classifications. 65N30, 65N55, 65N15.

1. Introduction. Let $R \times S^2 \times \mathcal{E}$ be the Cartesian product of a bounded domain $R \subset \Re^3$, the unit sphere S^2, and a bounded, non-negative interval $\mathcal{E}$. The time-independent Boltzmann transport equation is

$$[\Omega \cdot \nabla + \sigma_t(\mathbf{x}, E)]\, \psi(\mathbf{x}, \Omega, E) = \int dE' \int d\Omega' \sigma_s(\mathbf{x}', E' \to E, \Omega \cdot \Omega')\psi(\mathbf{x}', \Omega', E')$$
$$+\, q(\mathbf{x}, \Omega, E) \qquad (\mathbf{x}, \Omega, E) \in R \times S^2 \times \mathcal{E} \tag{1.1}$$

with boundary condition

$$\psi(\mathbf{x}, \Omega, E) = g(\mathbf{x}, \Omega, E) \qquad \mathbf{n} \cdot \Omega < 0,\ \mathbf{x} \in \partial R. \tag{1.2}$$

This equation models the transport of particles through an inhomogeneous medium. In particular, the Boltzmann equation is well known to model the transport of neutrons/photons. In this case, ψ is the angular flux, σ_t and σ_s are respectively the medium's total and scattering cross-sections ($\sigma_a := \sigma_t - \sigma_s$ is the absorption cross-section), q is the external source, and E is the energy of the particles. The integral source term describes the scattering of particles into different directions and energies.

Unfortunately, solving the linear Boltzmann equation is difficult: standard numerical schemes can be inaccurate and computationally inefficient. For example, the popular S_n angle discretization (collocation in angle, [14]) produces the so-called ray effects which pollute the numerical solution. This pollution can be viewed mathematically as "contamination"

* Received May 18, 2001. Accepted for publication September 12, 2001. Recommended by Joel Dendy.

†Center for Applied Scientific Computing, Lawrence Livermore National Laboratory, P.O. Box 808 L-561, Livermore, CA 94551. *email*: chang1@llnl.gov, lee123@llnl.gov. This work was performed under the auspices of the U.S. Department of Energy by Lawrence Livermore National Laboratory under contract no. W-7405-Eng-48.

from a poorly chosen finite element space for the angle component of the discretization– i.e., collocation in angle is equivalent to discretization with delta basis functions, which form a poor approximating basis set. Fortunately, a P_n discretization, which uses a better approximating basis set (i.e., spherical harmonics), eliminates these ray effects. But solving for the expansion coefficients or "lm moments" is difficult. The coefficients couple strongly with each other, creating a strongly coupled system of partial differential equations (PDE's); numerical algorithms for solving such strongly coupled systems are difficult to develop. In this paper, novel algorithms for solving this coupled system are presented. In particular, a multigrid algorithm for solving the P_n discretization of the linear Boltzmann equation using a first-order system least-squares (FOSLS) methodology ([17]) is described. The authors are unaware of any published work or existing codes that solve the full FOSLS P_n equations in a multilevel fashion.

This paper is an extension of the research reported in [9]. In that paper, a preconditioned conjugate gradient iteration with a block diagonal preconditioner was used to solve the system of P_n equations. Each block of this preconditioner describes only a single diagonal lm-lm coefficient coupling, but defined over the whole spatial domain rather than the full lm-$l'm'$ coupling. Thus, successively inverting each block of this preconditioner corresponds to successively solving only the lm-lm scalar PDE over the whole spatial domain. The numerical results presented in that paper confirm the non-scalibility of this algorithm with respect to both the number of spherical harmonic terms and the number of spatial nodes used in the full discretization. This non-scalibility reflects this scheme's inability to handle the strong intra- and inter-moment coupling.

In this paper, two algorithms that ameliorate some of the moment coupling are presented. One algorithm consists of a multigrid scheme for the spatial coupling of the P_n discretization. Here, the unknowns are updated moment-wise first and then spatial-wise. In this way, for a Gauss-Seidel relaxation scheme, at each spatial node in turn, every moment is updated before going to the next spatial node so that the full moment coupling is considered at a fixed node. Physically, since the Boltzmann equation describes the balancing of particle transfer, by solving for all the moments at a spatial node first, this relaxation somewhat enforces a local balancing of particle transfer at each spatial node. One may also use a preconditioned conjugate gradient iteration with a block diagonal preconditioner that describes the full l-l (i.e., moments lm-lm' with $-l \leq m, m' \leq l$) intra-moment coupling. Each of the diagonal blocks can be solved with a few cycles of this multigrid scheme restricted only to the l-l moment block. Comparing the convergence rates of this preconditioned conjugate gradient scheme and the above multigrid scheme will expose the relative strength of the intra- and inter-moment coupling in the P_n equations.

A system projection multilevel algorithm using the above Gauss-Seidel smoother performs well for parameter regimes that are thick and substantially absorptive. In the thick, low absorptive regime, or so-called region 3, the 1-1 moment system resembles a time-dependent incompressible elasticity equation. The problems with this latter equation are well known: e.g., possible finite element locking and problematic divergence-free near-nullspace components that impede standard multigrid performance ([3], [11], [21], [22]). For this incompressible equation, higher-order or Raviart-Thomas finite element spaces are often used to eliminate locking. These spaces also may induce discrete Helmholtz decompositions, which in turn, may implicitly improve multigrid performance ([3], [11], [21], [22]). But a closer look at the 1-1 system reveals that the corresponding system operator behaves more like $(I - \nabla\nabla\cdot)$. Locking even for low-order conforming finite elements now does not appear to be as severe as in the incompressible elasticity case. However, divergence-free error components are still problematic for standard multigrid solvers. These error components can be

ETNA
Kent State University
etna@mcs.kent.edu

highly oscillatory yet unaffected by standard nodal smoothers. But since these components are essentially local circulations ([7], [21]), effective smoothers exist. In particular, multiplicative/additive Schwarz smoothers that simultaneously update all unknowns in the support of these local circulations can effectively damp out these error components. We will use these smoothers in the multigrid solver.

This paper proceeds as follows. In section 2, a brief presentation of the notation and functional setting used throughout this paper is given. In section 3, a summary of the FOSLS theory developed in [17] for the mono-energetic, isotropic scattering Boltzman equation is reviewed. This theory shows that locally away from the material interfaces, by appropriately scaling the system of PDE's, the second-order moment coupling essentially describes the whole coupled system of PDE's. This fact will be used to develop our numerical schemes. In section 4, the $P_n - h$ finite element discretization for the FOSLS formulation is developed. The system of PDE's is explicitly described, and from this description, one can observe the difficulties in region 3. In section 5, a multigrid scheme is described for the mono-energetic, isotropic scattering case. Multigrid components (relaxation and coarse grid correction), and methods of homogenization of the fine grid material and scaling coefficients will be examined. In subsection 5.1, an algorithmic extension to the full multi-group, anisotropic scattering case will be given. Section 6 presents some numerical results. Computational scaling studies for the mono-energetic, isotropic scattering case will be presented for both the multigrid and preconditioned conjugate gradient schemes. For the multigrid scheme, these results show good scalibility with respect to the number of spatial nodes and linear growth with respect to the number of moments. For the preconditioned conjugate gradient scheme, these results show mild non-scalibility with respect to the number of spatial nodes and linear growth with respect to the number of moments. This difference signals a spatially smooth inter-moment coupling error mode that is not handled by the latter scheme. Section 6 also will present multigrid and discretization convergence results for region 3 problems and for a full multi-group, anisotropic scattering problem. These latter results, together with the analyis in Section 4, indicate that locking may not be severe for realistic Boltzmann transport problems.

2. Notation. We briefly present some of the notation and functional setting used throughout this paper. First, for any non-negative integer s, let $H^s(R)$ denote the usual Sobolev space of order s defined over R and with norm denoted by $\|\cdot\|_{s,R}$ (cf. [1]). For $s = 0$, the $L^2(R)$ is denoted by $\|\cdot\|_R$. When it is obvious that the norm is defined over R, subscript R will be omitted. Occasionally, we will have need of a general Hilbert space X. Its norm will be denoted by $\|\cdot\|_X$.

Further notations and definitions are needed for the FOSLS functional. This functional involves an L^2 term defined over $R \times S^2$. We denote this norm by $\|\cdot\|_{R,\Omega}$. This functional is also defined over the Sobolev space

$$V := \left\{ v \in L^2(R \times S^2) : \left(S^{-1}\Omega \cdot \nabla v, \Omega \cdot \nabla v\right)_{R,\Omega} + (Tv, v)_{R,\Omega} < \infty \right\}$$

with norms

$$\|v\|_V^2 := \left(S^{-1}\Omega \cdot \nabla v, \Omega \cdot \nabla v\right)_{R,\Omega} + (Tv, v)_{R,\Omega}$$

and

$$\|v\|_{V_1}^2 := \|v\|_V^2 + \int_{\partial R} \int_{\mathbf{n}\cdot\Omega<0} \psi\overline{\psi} |\mathbf{n} \cdot \Omega| \, d\Omega d\sigma.$$

Here, S^{-1} and T are linear operators that essentially act on the angular variable of a function $\psi(\mathbf{x}, \Omega)$.

To analyze the P_n equations, we will derive several relations involving the normalized spherical harmonics

$$Y_{lm} = \begin{cases} a_{lm}\, p_{l,m}(\cos\theta)\, \cos m\phi & m \geq 0 \\ a_{lm}\, p_{l,|m|}(\cos\theta)\, \sin |m|\phi & m < 0, \end{cases}$$

where

$$a_{lm} = \sqrt{\frac{(2l+1)(l-|m|)!}{2\pi(1+\delta_{m0})(l+|m|)!}}\,,$$

where θ and ϕ are related to Ω by $\Omega = (\sin\theta\cos\phi, \sin\theta\sin\phi, \cos\phi)$, and where $p_{l,m}$ is the lm'th associated Legendre polynomial ([15]). To notationally simplify these derivations, we will use the Dirac bra-ket notation ([20]). Considering the linear vector space generated by $\{Y_{lm}\}_{lm}$, vector element Y_{lm} will be denoted by

$$|lm> \qquad \text{the ket vector.}$$

Its dual vector is denoted by

$$< lm| \qquad \text{the bra vector.}$$

Given a linear operator A acting on ket $|lm>$, the $L^2(\Omega)$ inner product between $A|lm>$ and $< l'm'|$ is denoted

$$< l'm'|A|lm > .$$

Finally, since $\{Y_{lm}\}_{lm}$ forms a complete orthonormal set, we note that the completeness property

$$I = \sum_{lm} |lm >< lm|$$

holds.

3. Theory. For self-containment, we review some of the existing theory for the FOSLS formulation of (1.1)-(1.2). Except for the scaling operator for anisotropic scattering, which was communicated to us by T. Manteuffel ([18]), these results were derived in ([17]).

Theory for the FOSLS formulation of (1.1)-(1.2) has been developed only for the mono-energetic, isotropic scattering form of the Boltzmann equation. This simplified form of the Boltzmann equation is obtained by assuming the approximate energy separability of ψ and taking a truncated Legendre series expansion of σ_s ([14]):

(i) $\psi(\mathbf{x}, \Omega, E) \approx f(E)\psi^g(\mathbf{x}, \Omega)$, $E_g < E < E_{g+1}$, $g = 1, \cdots, N$, where f is a normalization function and

$$\psi^g(\mathbf{x}, \Omega) := \int_{E_g}^{E_{g+1}} dE\ \psi(\mathbf{x}, \Omega, E);$$

(ii)

$$\begin{aligned} \sigma_s(\mathbf{x}', E \to E', \Omega \cdot \Omega') &= \sum_{j=0}^{M} \sigma_{s,j}(\mathbf{x}') p_j(\Omega \cdot \Omega') \\ &= \sum_{j=0}^{M} \sigma_{s,j}(\mathbf{x}', E, E') \frac{1}{2j+1} \sum_{m=-j}^{j} Y_{jm}(\Omega)\overline{Y}_{jm}(\Omega'), \end{aligned}$$

ETNA
Kent State University
etna@mcs.kent.edu

where p_j is the j'th Legendre polynomial and Y_{jm} is the jm'th spherical harmonic ([14], [15], [20]).

A simple calculation then shows that (1.1) becomes a system PDE for the group fluxes $\psi^g(\mathbf{x},\Omega)$ with matrix operator

$$\begin{pmatrix} \left[\Omega\cdot\nabla+\sigma_t^{00}-\sum_{jm}\sigma_{s,j}^{0,0}P_{jm}\right] & -\sum_{jm}\sigma_{s,j}^{0,1}P_{jm} & \cdots & -\sum_{jm}\sigma_{s,j}^{0,N}P_{jm} \\ -\sum_{jm}\sigma_{s,j}^{1,0}P_{jm} & \left[\Omega\cdot\nabla+\sigma_t^{11}-\sum_{jm}\sigma_{s,j}^{1,1}P_{jm}\right] & \cdots & -\sum_{jm}\sigma_{s,j}^{1,N}P_{jm} \\ \vdots & \vdots & \ddots & \vdots \end{pmatrix}.$$

Here, the superscripts in the cross-section coefficients denote the $g-g'$ energy group coupling, and P_{jm} is the spherical harmonic projection onto Y_{jm}. In particlar, for a single energy group and isotropic scattering ($\sigma_{s,j}=0,\ j=1,\cdots,M$), supressing the super/subscripts, the boundary value problem is

$$\left\{\begin{array}{lll} [\Omega\cdot\nabla+\sigma_tI-\sigma_sP]\psi(\mathbf{x},\Omega)= & q & (\mathbf{x},\Omega)\in R\times S^2 \\ \psi(\mathbf{x},\Omega)= & g & \mathbf{x}\in\partial R,\ \mathbf{n}\cdot\Omega<0, \end{array}\right. \tag{3.1}$$

where the scattering term is

$$P\psi(\mathbf{x},\Omega)=\int_{S^2}\psi(\mathbf{x},\Omega)\,d\Omega.$$

For simplicity, we will assume that R is of unit diameter.

In the FOSLS formulation of (3.1), the Boltzmann operator is rewritten with the absorption cross-section

$$\begin{aligned} \mathcal{L}:&=\Omega\cdot\nabla+\sigma_t(I-P)+\sigma_aP \\ &=\Omega\cdot\nabla+T, \end{aligned} \tag{3.2}$$

where $T:=\sigma_t(I-P)+\sigma_aP$. Now introducing the scaling operator

$$S=\left\{\begin{array}{lll} I & \sigma_t\le 1 & \text{region 1: thin} \\ \sigma_t(I-P)+\sigma_aP & \sigma_t\ge 1 \text{ and } \sigma_a\ge\frac{1}{\sigma_t} & \text{region 2: thick with absorption} \\ \sigma_t(I-P)+\frac{1}{\sigma_t}P & \sigma_t\ge 1 \text{ and } \sigma_a\le\frac{1}{\sigma_t} & \text{region 3: thick with little absorption} \end{array}\right.$$

with inverse

$$\begin{aligned} S^{-1}&=\left\{\begin{array}{ll} (I-P)+P & \sigma_t\le 1 \\ \frac{1}{\sigma_t}(I-P)+\frac{1}{\sigma_a}P & \sigma_t\ge 1 \text{ and } \sigma_a\ge\frac{1}{\sigma_t} \\ \frac{1}{\sigma_t}(I-P)+\sigma_tP & \sigma_t\ge 1 \text{ and } \sigma_a\le\frac{1}{\sigma_t} \end{array}\right. \\ &=c_1(I-P)+c_2P, \end{aligned} \tag{3.3}$$

the space-angle FOSLS formulation is to minimize the scaled least-squares functional

$$\mathcal{F}(\psi;q,g):=\left\|S^{-\frac{1}{2}}(\mathcal{L}\psi-q)\right\|_{R,\Omega}^2+2\int_{\partial R}\int_{\mathbf{n}\cdot\Omega<0}(\psi-g)\overline{(\psi-g)}|\mathbf{n}\cdot\Omega|\,d\Omega d\sigma \tag{3.4}$$

over an appropriate Sobolev space. Note that because of the boundary integral in the least-squares functional, the inflow boundary condition need not be enforced on this Sobolev space.

ETNA
Kent State University
etna@mcs.kent.edu

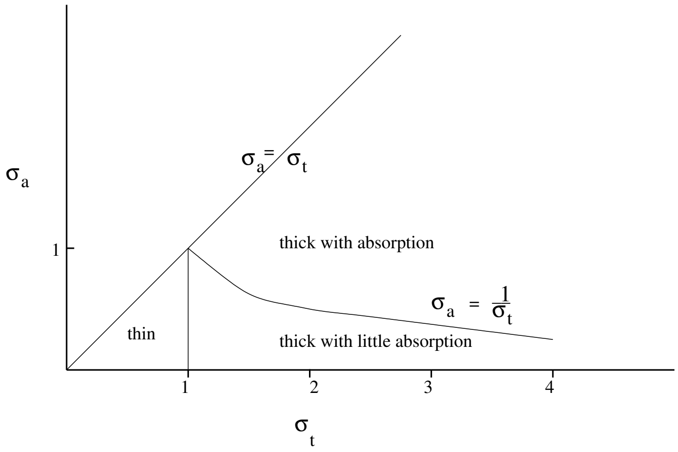

FIG. 3.1. *Parameter regimes defined by scaling operator S.*

The appropriate Sobolev space is V. It was shown in [17] that $\mathcal{F}$ is equivalent to the V_1-norm over V. Thus, functional

$$\mathcal{F}(\psi;0,0) = \left(S^{-1}\Omega\cdot\nabla\psi,\Omega\cdot\nabla\psi\right)_{R,\Omega} + \left(S^{-1}T\psi,\Omega\cdot\nabla\psi\right)_{R,\Omega} + \left(S^{-1}\Omega\cdot\nabla\psi,T\psi\right)_{R,\Omega}$$
$$+ (T\psi,T\psi)_{R,\Omega} + 2\int_{\partial R}\int_{\mathbf{n}\cdot\Omega<0}\psi\overline{\psi}|\mathbf{n}\cdot\Omega|\,d\Omega d\sigma$$

is equivalent to

$$\|\psi\|^2_{V_1} = \left(S^{-1}\Omega\cdot\nabla\psi,\Omega\cdot\nabla\psi\right)_{R,\Omega} + (T\psi,\psi)_{R,\Omega} + \int_{\partial R}\int_{\mathbf{n}\cdot\Omega<0}\psi\overline{\psi}|\mathbf{n}\cdot\Omega|\,d\Omega d\sigma.$$

That is, the first-order terms $\left(S^{-1}T\psi,\Omega\cdot\nabla\psi\right)_{R,\Omega}$ and $\left(S^{-1}\Omega\cdot\nabla\psi,T\psi\right)_{R,\Omega}$ are majorized by the second-order term $\left(S^{-1}\Omega\cdot\nabla\psi,\Omega\cdot\nabla\psi\right)_{R,\Omega}$.

Minimizing $\mathcal{F}$ over V is effectively solving the variational equation

$$a(\psi,w) := \left(S^{-1}\mathcal{L}\psi,\mathcal{L}w\right)_{R,\Omega} + 2\int_{\partial R}\int_{\mathbf{n}\cdot\Omega<0}\psi\overline{w}|\mathbf{n}\cdot\Omega|\,d\Omega d\sigma$$
$$= (q,S^{-1}\mathcal{L}w)_{R,\Omega} + 2\int_{\partial R}\int_{\mathbf{n}\cdot\Omega<0}g\overline{w}|\mathbf{n}\cdot\Omega|\,d\Omega d\sigma \tag{3.5}$$

for all $v\in V$. Because of the norm equivalence, one essentially needs to develop an effective solver or preconditioner for the discrete system corresponding to the bilinear form

$$b(v,w) := \left(S^{-1}\Omega\cdot\nabla v,\Omega\cdot\nabla w\right)_{R,\Omega} + (Tv,w)_{R,\Omega} + \int_{\partial R}\int_{\mathbf{n}\cdot\Omega<0}v\overline{w}|\mathbf{n}\cdot\Omega|\,d\Omega d\sigma.$$

A scalable solution method for the minimization of the least-squares functional will require a scalable solver for this latter system.

For the multi-group, anisotropic scattering case, the FOSLS formulation is generalized by using the scaling operator

$$S^{-\frac{1}{2}} = c_1 I + \sum_{j=0}^{M} c_j P_j$$

ETNA
Kent State University
etna@mcs.kent.edu

$$= \begin{cases} I & \sigma_t \le 1 \\ \frac{1}{\sqrt{\sigma_t}} I + \sqrt{\frac{\sigma_{s,j}}{\sigma_t(\sigma_t - \sigma_{s,j})}}\, P_j & \sigma_t > 1 \text{ and } (\sigma_t - \sigma_{s,0}) \ge \frac{1}{\sigma_t} \\ \frac{1}{\sqrt{\sigma_t}} I + \sqrt{\frac{\left(\sigma_t - \frac{1}{\sigma_t}\right)\frac{\sigma_{s,j}}{\sigma_{s,0}}}{\sigma_t\left[\sigma_t + \left(\sigma_t - \frac{1}{\sigma_t}\right)\frac{\sigma_{s,j}}{\sigma_{s,0}}\right]}}\, P_j & \sigma_t > 1 \text{ and } (\sigma_t - \sigma_{s,0}) < \frac{1}{\sigma_t}. \end{cases}$$

in (3.4) ([18]).

4. Spherical Harmonic-h ($\mathbf{P_n} - \mathbf{h}$) Finite Element Discretization. Two of the advantages of a FOSLS formulation are that it leads to symmetric positive-definite linear systems, and that it is endowed with a computable a posteriori error measure ([8]). For the Boltzmann equation, the symmetric positive-definiteness allows such efficient linear system solvers as multigrid and preconditioned conjugate gradient schemes to be used on the P_n discretization of the FOSLS variational form, and the a posteriori error measure leads to good, simple local grid refinement strategies. Indeed, standard Galerkin P_n discretizations of the Boltzmann equation lead to non-symmetric linear systems that are difficult to solve efficiently, and standard Galerkin P_n discretizations do not naturally lead to simple computable error measures. There are efficient Petrov-Galerkin formulations of the S_n discretization, but this discretization suffers from the ray effect in the thin region and in thick regions when the source is close to the points of observation (e.g., points along the boundary). Nevertheless, the P_n FOSLS method is not immuned from problems itself, as will be shown later. But even with these difficulties, the P_n FOSLS method is a scheme that can handle all parameter regimes, with the above attractive FOSLS features.

The P_n discretization consists of taking a truncated spherical harmonic expansion of the angular flux:

$$\begin{aligned} \psi(\mathbf{x}, \Omega) &\approx \psi_N(\mathbf{x}, \Omega) \\ &= \sum_{l=0}^{N} \sum_{m=-l}^{l} \phi_{lm}(\mathbf{x}) Y_{lm}(\Omega). \end{aligned} \tag{4.1}$$

The ϕ_{lm}'s are the moments or generalized Fourier coefficients. We will consider only the mono-energetic, isotropic scattering problem. Substituting ψ_N into bilinear form $a(\cdot,\cdot)$, and testing it against $v(\mathbf{x})Y_{l'm'}(\Omega)$, $l' = 0, ..., N$ and $m' = -l', ..., l'$, a semi-discretization is obtained. Now because P acts only on the angular variable, we have

$$[I - P]\phi_{l,m}(\mathbf{x})Y_{lm}(\Omega) = \phi_{l,m}(\mathbf{x})[(I - P)Y_{lm}(\Omega)],$$

and so, T and S^{-1} simply project the zero and non-zero moments differently. Moreover, because of the V_1 norm equivalence, to analyze this semi-discrete system, only the zeroth-order and second-order terms need to be considered.

For the zeroth-order term, since T acts only on the angular variable, we have

$$(T\psi_N, vY_{l',m'})_{R,\Omega} = \sum_{lm} < Y_{lm}|T|Y_{lm} > (\phi_{lm}, v)_R\,, \tag{4.2}$$

where $< \cdot|A|\cdot >$ is the bra-ket notation for the angular inner product with operator A acting on ket $|\cdot >$, and where $(\cdot,\cdot)_R$ is the spatial inner product over R. For the second-order term, we have

$$\left(S^{-1}\Omega\cdot\nabla\psi_N, \Omega\cdot\nabla v Y_{l',m'}\right)_{R,\Omega} = \sum_{i=1}^{3}\sum_{j=1}^{3}\sum_{lm} \left(S^{-1}\Omega_i Y_{lm}\phi_{lm,i}, \Omega_j Y_{l'm'} v_j\right)_{R,\Omega}$$

$$= \sum_{i=1}^{3}\sum_{j=1}^{3}\sum_{lm} < Y_{lm}|\Omega_i S^{-1}\Omega_j|Y_{l'm'} > (\phi_{lm,i}, v_j)_R . \tag{4.3}$$

Here, i and j denote spatial differentiation. Note that the sparsity pattern of the second-order stiffness matrix depends on both the spatial differentiation operators and the moment coupling created through

$$< Y_{lm}|\Omega_i S^{-1}\Omega_j|Y_{l'm'} > .$$

Consider the diagonal lm-lm element of $< Y_{lm}|\Omega_i S^{-1}\Omega_j|Y_{l'm'} >$. This element can be viewed as a full 3×3 tensor describing the "diffusive" interaction of moment ϕ_{lm} with itself. Viewed this way, $< Y_{lm}|\Omega_i S^{-1}\Omega_j|Y_{l'm'} >$ is a $(N+1)^2 \times (N+1)^2$ block matrix of 3×3 tensors with each lm-$l'm'$ tensor describing the spatial anisotropy coupling between ϕ_{lm} and $\phi_{l'm'}$. Fortunately, this block matrix of tensor has some structure. To see this, the completeness property

$$\sum_{l''m''} |Y_{l''m''} >< Y_{l''m''}| = I$$

of spherical harmonics ([20]) is needed. Applying this identity twice, we have

$$\begin{aligned} < Y_{lm}|\Omega_i S^{-1}\Omega_j|Y_{l'm'} > &= \sum_{l''m''} < Y_{lm}|\Omega_i S^{-1}|Y_{l''m''} >< Y_{l''m''}|\Omega_j|Y_{l'm'} > \\ &= \sum_{l''m''}\sum_{l'''m'''} < Y_{lm}|\Omega_i|Y_{l'''m'''} > \\ &\qquad < Y_{l'''m'''}|S^{-1}|Y_{l''m''} >< Y_{l''m''}|\Omega_i|Y_{l'm'} > . \end{aligned}$$

But S^{-1} simply scales the ket $|Y_{l''m''} >$ by

$$\begin{cases} c_1 & l'' \neq 0 \\ c_2 & l'' = 0 \end{cases}$$

(cf., equation (3.3)). The orthogonality of spherical harmonics then implies

$$< Y_{l'''m'''}|S^{-1}|Y_{l''m''} >= \begin{cases} c_1\ \delta_{l''l'''}\delta_{m''m'''} & l'' \neq 0 \\ c_2\ \delta_{l''0}\delta_{m''0} & l'' = 0, \end{cases}$$

and so,

$$< Y_{lm}|\Omega_i S^{-1}\Omega_j|Y_{l'm'} > = \sum_{l''m''} c_{l''m''} < Y_{lm}|\Omega_i|Y_{l''m''} >< Y_{l''m''}|\Omega_j|Y_{l'm'} >, \tag{4.4}$$

where $c_{l''m''} = c_1$ if $l \neq 0$ and $c_{00} = c_2$. Moreover, it can be shown that

$$< Y_{\tilde{l}\tilde{m}}|\Omega_i|Y_{\hat{l}\hat{m}} >\neq 0$$

only when $\hat{l} = \tilde{l} \pm 1$. Thus, $< Y_{lm}|\Omega_i S^{-1}\Omega_j|Y_{l'm'} >$ is a weighted product of two block tridiagonal matrices, which implies that it is block pentadiagonal. In fact, further properties of spherical harmonics show that this block pentadiagonal matrix is nonzero only when $l = l' \pm 2$. Hence, the even and odd moments decouple in the second-order term. Figure 4.1 illustrates this structure for $N = 4$, where the '$\times$' blocks are the non-zero block entries and the small diagonal blocks are the lm-lm 3×3 tensors. Finer structure of this block

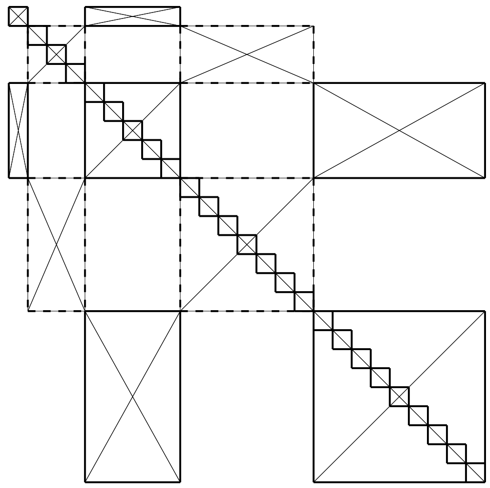

FIG. 4.1. *Intra-moment coupling structure with diagonal lm-lm blocks.*

pentadiagonal matrix can be found by using additional properties of the spherical harmonics (with respect to m).

Now, assume a finite element discretization of the spatial component. Using the test function $b_\beta(\mathbf{x})Y_{l'm'}(\Omega)$, where $\{b_\beta\}_\beta$ is a basis set for the spatial finite element space, the second-order term becomes

$$\sum_{i=1}^{3}\sum_{j=1}^{3}\sum_{lm}\sum_{\alpha} < Y_{lm}|\Omega_i S^{-1}\Omega_j|Y_{l'm'} > (b_{\alpha,i}, b_{\beta,j})_R\ \phi_{lm,\alpha} =$$

$$\sum_{i=1}^{3}\sum_{j=1}^{3}\sum_{lm}\sum_{\alpha}\left[\sum_{l''m''} c_{l''m''} < Y_{lm}|\Omega_i|Y_{l''m''} >< Y_{l''m''}|\Omega_j|Y_{l'm'} >\right] \times (b_{\alpha,i}, b_{\beta,j})_R\ \phi_{lm,\alpha}. \tag{4.5}$$

Here, if b_α at spatial node α is the standard hat function, then $\phi_{lm,\alpha}$ is the value of the lm moment at that node. Using the structure of $< Y_{lm}|\Omega_i S^{-1}\Omega_j|Y_{l'm'} >$, the structure of the full discrete second-order term is also block pentadiagonal with total size $M(N+1)^2 \times M(N+1)^2$ where M is the number of spatial nodes. Corresponding to each 3×3 tensor of $< Y_{lm}|\Omega_i S^{-1}\Omega_j|Y_{l'm'} >$ is an $M\times M$ submatrix describing the discretized spatial coupling of moment lm to $l'm'$. Alternatively, assuming R to be tessalated into tetrahedral elements and assuming $\{b_\alpha\}$ to be trilinear finite elements, the second-order term can be re-ordered to have a 27 block stripe structure corresponding to the 27 point stencil of the spatial differentiation operator. Each block on any stripe gives the $< Y_{lm}|\Omega_i S^{-1}\Omega_j|Y_{l'm'} >$ coupling at a spatial point. Such ordering is better for computation.

Since bilinear form $b(\cdot,\cdot)$ also contains (4.3), an effective solver or preconditioner must be able to efficiently invert this complicated linear system. However, for some parameter regimes, a more sophisticated spatial discretization and a non-standard multigrid scheme may be needed. A problem arises because scaling coefficients c_1 and c_2 may differ by orders of magnitude. Thus, on one hand, the scaling operator leads to the correct asymptotic limiting equation in these regimes ([16], [17]), but on the other, a complicated discretization and multigrid scheme may be required.

To see the scaling problem, from (4.4), we see that only the $l''$$-$$m''$ =0-0 column and row of $< Y_{lm}|\Omega_i|Y_{l''m''} >$ and $< Y_{l''m''}|\Omega_j|Y_{l'm'} >$, respectively, are scaled by $\sqrt{c_2}$. All other rows and columns are scaled by $\sqrt{c_1}$. Because $< Y_{lm}|\Omega_i|Y_{l''m''} >$ and $< Y_{l''m''}|\Omega_j|Y_{l'm'} >$ are block tridiagonal and because $< Y_{lm}|\Omega_i S^{-1}\Omega_j|Y_{l'm'} >$ is nonzero only when $l' = l \pm 2$, then only diagonal moment blocks l-l =0-0 and l-l =1-1 of the second-order term can contain c_2 scaled terms. In particular, for regions 1, 2, and 3 respectively, these blocks are

$$\begin{bmatrix} -\frac{1}{3}\nabla\cdot\nabla & 0 \\ 0 & -\frac{1}{5}\underline{\Delta} - \frac{6}{15}\nabla\nabla\cdot \end{bmatrix} \qquad c_1 = 1, c_2 = 1,$$

$$\begin{bmatrix} -\frac{1}{3\sigma_t}\nabla\cdot\nabla & 0 \\ 0 & -\frac{1}{5\sigma_t}\underline{\Delta} - \left(\frac{1}{3\sigma_a} + \frac{1}{15\sigma_t}\right)\nabla\nabla\cdot \end{bmatrix} \qquad c_1 = \frac{1}{\sigma_t}, c_2 = \frac{1}{\sigma_a},$$

and

$$\begin{bmatrix} -\frac{1}{3\sigma_t}\nabla\cdot\nabla & 0 \\ 0 & -\frac{1}{5\sigma_t}\underline{\Delta} - \left(\frac{\sigma_t}{3} + \frac{1}{15\sigma_t}\right)\nabla\nabla\cdot \end{bmatrix} \qquad c_1 = \frac{1}{\sigma_t}, c_2 = \sigma_t$$

([17]). Block l-l =0-0 is Laplacian, so poses no problems. However, block l-l =1-1 contains the grad-div operator $\nabla\nabla\cdot$, whose nullspace consists of divergence-free functions. In particular, in region 3 and region 2 when $\sigma_t \gg 1$ and $\sigma_a \approx \frac{1}{\sigma_t}$, since the 1-1 block is majorized by the grad-div operator, divergence-free components create problems for standard iterative solvers.

These problems remain even when both the zeroth and second-order terms of $b(\cdot,\cdot)$ are considered. Now the diagonal blocks in regions 2 and 3 are

$$\begin{bmatrix} \sigma_a I - \frac{1}{3\sigma_t}\nabla\cdot\nabla & 0 \\ 0 & \sigma_t I - \frac{1}{5\sigma_t}\underline{\Delta} - \left(\frac{1}{3\sigma_a} + \frac{1}{15\sigma_t}\right)\nabla\nabla\cdot \end{bmatrix}, \tag{4.6}$$

and

$$\begin{bmatrix} \sigma_a^2\sigma_t I - \frac{1}{3\sigma_t}\nabla\cdot\nabla & 0 \\ 0 & \sigma_t I - \frac{1}{5\sigma_t}\underline{\Delta} - \left(\frac{\sigma_t}{3} + \frac{1}{15\sigma_t}\right)\nabla\nabla\cdot \end{bmatrix}. \tag{4.7}$$

These divergence-free components are approximate eigenfunctions of the l-l =1-1 block corresponding to eigenvalue σ_t.

To further expose the difficulties of the 1-1 block, one can compare it to a semi-discrete form of a time-dependent incompressible elasticity equation:

$$\begin{cases} [cI - \underline{\Delta} - \lambda\nabla\nabla\cdot]\,\mathbf{u} & = \mathbf{f} \quad \text{in } R \\ \mathbf{u} & = \mathbf{0} \quad \text{on } \partial R \end{cases} \tag{4.8}$$

with $c = O(1/\Delta t)$ and $\lambda \gg 1$. Not only do divergence-free error components complicate the system solver, but also the locking effect degrades uniform discretization convergence with respect to λ ([4], [5], [6]). Assuming conforming piecewise linear finite elements and full H^2-regularity, this non-uniformity can be anticipated from the usual error estimate

$$\|\mathbf{u} - \mathbf{u}^h\|_1 \le C(\lambda) h \|\mathbf{u}\|_2 \tag{4.9}$$

with constant $C(\lambda)$ dependent on λ. But to examine locking more precisely, let a_0 be a continuous, coercive bilinear form defined over a Hilbert space X (i.e., $\alpha\|v\|_X^2 \le a_0(v,v) \le \beta\|v\|_X^2\ \forall v \in X$), let $B : X \to L^2$ be a continuous mapping, and consider the weak equation

$$a_0(u_\lambda, v) + \lambda(Bu_\lambda, Bv) = (f, v) \qquad \forall v \in X.$$

ETNA
Kent State University
etna@mcs.kent.edu

Locking occurs over the finite element space X^h if

$$X^h \cap N(B) = \{0\} \tag{4.10}$$

and if

$$\|v_h\|_X \leq C(h)\|Bv^h\| \qquad \forall v^h \in X^h. \tag{4.11}$$

For (4.8), $X = H^2(R) \cap H_0^1(R)$,

$$a_0(\mathbf{u}_\lambda, \mathbf{v}) = (\nabla \mathbf{u}_\lambda, \nabla \mathbf{v}) + (c\mathbf{u}, \mathbf{v}),$$

and $B = \nabla\cdot$. Conditions (4.10)-(4.11) guarantee that for any fixed h and sufficiently large λ, there is a $\mathbf{v} \in H^2(R) \cap H_0^1(R)$ satisfying

$$[cI - \underline{\Delta} - \lambda\nabla\nabla\cdot]\,\mathbf{v} = \mathbf{f_v}$$

such that the relative error is bounded below by a constant independent of h :

$$C \leq \frac{|\mathbf{v} - \mathbf{v}^h|_1}{\|\mathbf{f_v}\|}. \tag{4.12}$$

Following the technique in [6], one can show that $\mathbf{v}$ is divergence-free with $\|\mathbf{v}\|_1 > 0$. Hence, since the solution of (4.8) becomes more incompressible as a function of λ, discretization convergence will not be uniform in λ.

Note that for a divergence-free function $\mathbf{v}$ with nonzero H^1-norm, $\mathbf{f_v}$ is nonzero and independent of λ. Using (4.9) and (4.12), we have

$$C \leq \frac{|\mathbf{v} - \mathbf{v}^h|_1}{\|\mathbf{f_v}\|} \leq \frac{C_2 C(\lambda) h \|\mathbf{v}\|_2}{\|\mathbf{f_v}\|}, \tag{4.13}$$

where C_2 independent of λ. From (4.13), we see that the severity of locking can be reduced if $\|\mathbf{f_v}\|$ depends similarly on λ as $C(\lambda)$ does.

Now consider the 1-1 block in (4.7) or (4.6) when $\sigma_t \approx \frac{1}{\sigma_a} \gg 1$. In either case, one has the approximate form

$$\begin{aligned}\left[c_1\sigma_t^2 I - \underline{\Delta} - c_2\sigma_t^2\nabla\nabla\cdot\right]\mathbf{u} &= \left[\sigma_t^2(c_1 I - c_2\nabla\nabla\cdot) - \underline{\Delta}\right]\mathbf{u} \\ &= \sigma_t \mathbf{f} := \mathbf{f_u}.\end{aligned} \tag{4.14}$$

Since a FOSLS P_n formulation of (3.1) leads to a "displacement" formulation of (4.14), $a_0(\mathbf{u}, \mathbf{v}) = (\nabla\mathbf{u}, \nabla\mathbf{v})$ and $B\mathbf{u} = (c_1\mathbf{u}, c_2\nabla\cdot\mathbf{u})^t$:

$$(\nabla\mathbf{u}, \nabla\mathbf{v}) + \sigma_t^2\left[(c_1\mathbf{u}, \mathbf{v}) + (c_2\nabla\cdot\mathbf{u}, \nabla\cdot\mathbf{v})\right] = (\mathbf{f_u}, \mathbf{v}).$$

Also, using linear finite elements, we have (4.9) with

$$C(\sigma_t^2) \approx \sigma_t^2. \tag{4.15}$$

Clearly, $(B\mathbf{u}, B\mathbf{v})$ corresponds to a scaled $H(\mathrm{div})$ norm of $\mathbf{u}$, and hence $N(B) = \{\mathbf{0}\}$, and consequently $\mathbf{f_u}$ must depend on σ_t^2. Indeed, from (3.5), the righthand-side associated with the 1-1 block is

$$\mathbf{f} = \frac{\sigma_t}{3}\nabla q_{00} + \text{lower order } \sigma_t \text{ terms},$$

where q_{00} is the zero'th moment of the external source q. For neutronic problems, $\|\nabla q_{00}\| = O(1)$ ([13], [16]) in the region 3. Thus,

$$\mathbf{f_u} = C_3 \sigma_t^2 \tilde{\mathbf{f}}. \tag{4.16}$$

Substituting (4.15)-(4.16) into (4.13), we have

$$\begin{aligned} \frac{|\mathbf{u} - \mathbf{u}^h|_1}{\|\mathbf{f_u}\|} &\leq \frac{C_2 \sigma_t^2 h \|\mathbf{u}\|_2}{C_3 \sigma_t^2 \|\tilde{\mathbf{f}}\|} \\ &= \frac{C_4 h \|\mathbf{u}\|_2}{\|\tilde{\mathbf{f}}\|}. \end{aligned} \tag{4.17}$$

But, even though (4.14) and (4.16) imply that $\mathbf{u}$ approximately satisfies

$$[c_1 I - c_2 \nabla\nabla\cdot]\mathbf{u} = C_3 \tilde{\mathbf{f}},$$

the upper bound in (4.17) does not imply uniform convergence. Consider the case when $\mathbf{u} = (u^1, u^2, u^3)$ satisifies any of the following conditions:

$$\begin{cases} u^2_{yx} + u^3_{zx} & = O(\sigma_t^2) \\ u^1_{xy} + u^3_{zy} & = O(\sigma_t^2) \\ u^1_{xz} + u^2_{yz} & = O(\sigma_t^2). \end{cases}$$

For this $\mathbf{u}$, $\|\mathbf{u}\|_2$ can be $O(\sigma_t^2)$.

5. A Multigrid Algorithm. The solution procedure involves minimizing $\mathcal{F}$ over an appropriate subspace of V. To accomplish this, a Rayleigh-Ritz finite element method is employed in the spatial discretization. Let $\mathcal{T}_h$ be a triangulation of domain R into elements of maximal length $h = \max\{\text{diam}(K) : K \in \mathcal{T}_h\}$, and let V^h be a finite dimensional subspace of V having the approximation property

$$\inf_{v^h \in V^h} \|v - v^h\|_{1,R} \leq Ch\|v\|_{2,R}$$

for all $v \in \left[H^2(R) \times L^2(S^2)\right]$. The $P_n - h$ finite element space is then

$$V_N^h := \left\{ v_N^h \in V^h : v_N^h = \sum_{l=0}^{N} \sum_{m=-l}^{l} \phi_{lm}^h(x) Y_{lm}(\Omega) \right\}.$$

The discrete fine grid minimization problem is

- Find $\psi_N^h \in V_N^h$ such that

$$\mathcal{F}(\psi_N^h; q, g) = \min_{v_N^h \in V_N^h} \mathcal{F}(v_N^h; q, g).$$

Equivalently, the discrete problem is

- Find $\psi_N^h \in V_N^h$ such that

$$a(\psi_N^h, v_N^h) = \left(q, S^{-1}\mathcal{L}v_N^h\right)_{R,\Omega} + 2\int_{\partial R}\int_{\mathbf{n}\cdot\Omega<0} g\overline{v}_N^h |\mathbf{n}\cdot\Omega|\, d\Omega d\sigma$$

for all $v_N^h \in V_N^h$.

(Although one actually solves for the moments ϕ^h_{lm}, we will not make this notational distinction in the algorithm.)

A standard projection multilevel scheme for solving either discrete problem is fairly straightforward. Let

$$\mathcal{T}_{2^{m-1}h} \subset \mathcal{T}_{2^{m-2}h} \subset \cdots \subset \mathcal{T}_{2h} \subset \mathcal{T}_h$$

be a conforming sequence of coarsenings of triangulation $\mathcal{T}_h$, let

$$V^m_N \subset V^{m-1}_N \subset \cdots \subset V^2_N \subset V^1_N := V^h_N$$

be a set of nested coarse grid subspaces of V^1_N, the finest subspace, and let

$$B^j = \left\{ b^j_{\nu,lm} \right\}$$

be a suitable (generally local in space) basis set for V^j_N. (For example, $b^j_{\nu,lm} = b^j_\nu Y_{lm}$, where b^j_ν is the standard piecewise linear hat function.) Given an initial approximation ψ^h_N on level j, the level j relaxation sweep consists of the following cycle

- for each $\nu = 1, 2, ..., M_j$, (M_j being the number of spatial nodes on grid j)

  for each lm, $0 \le l \le N,\ -l \le m \le l$,

$$\psi^h_N \leftarrow \psi^h_N + \alpha b^j_{\nu,lm},$$

  where α is chosen to minimize

$$\mathcal{F}\left(\psi^h_N + \alpha b^j_{\nu,lm}; q, g\right). \tag{5.1}$$

Since $\mathcal{F}\left(\psi^h_N + \alpha b^j_{\nu,lm}; q, g\right)$ is a quadratic function in α, this local minimization process is simple, and is, in fact, a Gauss-Seidel iteration. Moreover, note that the loops range over the moments first so that all moments are updated at a fixed spatial node before going onto the next spatial node. Note also that the search direction may involve more than one element of B^j. For such subspace iteration, denoting the direction by $\mathbf{b}^j_\nu$, one then needs to find $\boldsymbol{\alpha}$ that minimizes

$$\mathcal{F}\left(\psi^h_N + \boldsymbol{\alpha}\mathbf{b}^j_\nu; q, g\right).$$

A good choice for $\mathbf{b}^j_\nu$ is the subset

$$\{b^j_\nu Y_{lm} :\ 0 \le l \le N,\ -l \le m \le l,\ \nu \text{ fixed}\}$$

which results in a block Gauss-Seidel iteration that simultaneously updates all the moments at node ν.

Now, given a fine grid approximation ψ^h_N on level 1, the level 2 coarse grid problem is to find a correction ψ^{2h}_N such that

$$\mathcal{F}\left(\psi^h_N + \psi^{2h}_N; q, g\right) = \min_{v^{2h}_N \in V^2_N} \mathcal{F}\left(\psi^h_N + v^{2h}_N; q, g\right).$$

Having obtained this correction, ψ^h_N is updated according to

$$\psi^h_N \leftarrow \psi^h_N + \psi^{2h}_N.$$

ETNA
Kent State University
etna@mcs.kent.edu

Applying this procedure recursively yields a multilevel scheme in the usual way.

Away from the thick diffusive regime and for homogeneous material, this multigrid algorithm has the usual optimal multigrid efficiency. But for inhomogeneous materials even in the thin diffusive regime, when the material and scaling coefficients have fine-scale structure, the computational efficiency of this algorithm degrades; i.e., to preserve these fine-scale structures on the coarse grid, fine-scale computation is needed on the coarser levels. For a matrix-free implementation, the total cost of this fine-scale coarse grid computation is not nominal. For this reason, these coefficients should be homogenized to coarse-scale resolution and then used on the coarse grid. Coarse grid calculations now can be performed with coarse-scale computation.

Assuming that the coefficient jumps are grid-aligned on the finest grid, a simple homogenization method that can be applied is an averaging process. For example, the material and scaling coefficients can be arithmetically and harmonically averaged, respectively ([2], [10], [19], [23]): if $\{c_{lm,\mu_j^h}\}$, $\{\sigma_{t,\mu_j^h}\}$, and $\{\sigma_{a,\mu_j^h}\}$ are the coefficients in disjoint elements $\{\mu_j^h\}_{j=1}^r$ on grid h, and if the coarse grid element γ^{2h} is composed of the agglomerate

$$\gamma^{2h} = \cup_{j=1}^r \mu_j^h,$$

then

$$c_{lm,\gamma^{2h}} = \frac{r}{\sum_{j=1}^r \frac{1}{c_{lm,\mu_j^h}}} \tag{5.2}$$

$$\sigma_{t,\gamma^{2h}} = \frac{1}{r}\sum_{j=1}^r \sigma_{t,\mu_j^h} \tag{5.3}$$

$$\sigma_{a,\gamma^{2h}} = \frac{1}{r}\sum_{j=1}^r \sigma_{a,\mu_j^h}. \tag{5.4}$$

Here, harmonic averaging is more suitable for the scaling coefficients because they contribute to the diffusion tensors ([10]).

But one should note that using homogenized coefficients on coarser grids leads to a violation of a projection multilevel principle. Each coarse grid problem corresponds to a different minimization problem (i.e., non-nested bilinear forms). Thus, a coarse grid correction now is not an optimal subspace correction to the fine grid problem. In particular, for rapidly varying coefficients, it is possible for a coarse grid correction from a very coarse level to completely pollute the fine grid solution. Thus, more sophisticated homogenization schemes may be needed for complicated physics.

A sophisticated technique is also needed for region 3. Recall that the algebraically smooth error in this regime is predominantly controlled by the divergence-free error components of the 1-1 block. Since these components can be geometrically oscillatory, the smoother must eliminate the highly oscillatory ones. But since divergence-free functions are intrinsically vector quantities that are represented well only over the union of several elements, a point/nodal smoother is not sufficient. Rather, following [3], [11], and [21], an additive/multiplicative Schwarz smoother should be used. This block smoother must resolve the smallest representable circulation, or local divergence-free functions. Hence, the directions $\mathbf{b}_{\boldsymbol{\nu}}^j$ for this smoother have small spatial suppport so that the block solvers in a relaxation sweep involve only small linear systems for *only* the 1-1 block.

We now have a well-defined multigrid algorithm for all parameter regimes and for inhomogeneous material. Using an appropriate direction in the relaxation, one may handle the local intra- and inter-moment coupling well. One can also restrict this multigrid algorithm to the l-l blocks to produce a block preconditioner. Unlike the preconditioner described in [9], this preconditioner considers the full intra-moment coupling. By comparing the performances of this preconditioned conjugate gradient method and the multigrid algorithm defined over all lm-$l'm'$ coupling, one may be able to deduce the nature of the intra- and inter-moment coupling in the P_n equations.

5.1. An Algorithm for the Multi-Group, Anisotropic Scattering Equation. In section 3, we saw that (1.1) can be semi-discretized into a block system with mono-energetic, anisotropic scattering equations along the diagonal. Each of these diagonal equations can be solved with either the multigrid or preconditioned conjugate gradient algorithms described in the previous section. Thus, one obtains a scheme for solving the full multi-group, anisotropic scattering Boltzmann equation by using an outer block Gauss-Seidel iteration over the groups. For the important down-scattering problems in photon/neutron applications (i.e., particles can be scattered only into lower energy groups), this block iteration becomes a back solve. For general scattering, since the differential operators majorize the integral projection operators, the block system is diagonally dominant, and so, this block Gauss-Seidel iteration should perform well.

6. Numerical Experiments. The above $P_n - h$ finite element discretization of the FOSLS formulation was implemented. Angle integrals involving spherical harmonics were computed using analytical formulas, and the spatial moments were discretized with piecewise trilinear functions on rectangular solids. We conducted three sets of experiments. The first set examines the scalability of our code for region 1 and 2 problems. We consider both scalability with respect to the spatial meshsize, and scalability with respect to the number of processors used. The second set of experiments examines region 3 problems. Since our code does not have a parallel implementation of our multiplicative Schwarz smoother, only scalability with respect to the spatial meshsize is considered. However, we also examine discretization convergence to illustrate locking-free error for region 3 problems. Finally, the third set of experiments examines a full multi-group, anisotropic scattering problem. Both multigrid convergence rates and discretization error are considered.

i) Regions 1 and 2 Scaling Studies. The goal in these experiments is to investigate scalability with respect to the number of spatial nodes and spherical harmonics, and scalability with respect to the number of processors. Only the mono-energetic, isotropic scattering equation was considered since solving the diagonal equations in the semi-discrete multi-group, anisotropic scattering problem is the major task in the block Gauss-Seidel iteration described in Section 5.1. Also, only homogeneous material was considered. Results for the test suite Kobayashi problems ([12]) involving inhomogeneous material with large jumps will be presented in a future paper. Convergence rates for these benchmark problems are very similar to the rates presented here. For the current experiments, the source terms were taken to be zero and the initial guess was random for all moments. $V(1,1)$ cycles with "nodal moment" Gauss-Seidel relaxation were used in the multigrid and preconditioned conjugate gradient schemes. For the latter scheme, only one cycle was performed in the preconditioning solve. The stopping criterion was

$$\frac{\sqrt{\mathcal{F}(\psi^n;0,0)} - \sqrt{\mathcal{F}(\psi^{n-1};0,0)}}{\sqrt{\mathcal{F}(\psi^0;0,0)}} < 10^{-7}.$$

Results are tabulated in Tables 6.1 and 6.2.

Region	N	h	# Procs	# Unks/Proc	Iter	Time
1	1	1/32	1	$143,748$	9	279
	1	1/64	8	$137,313$	9	261
	1	1/128	64	$134,168$	9	267
	1	1/256	512	$132,614$	9	270
	3	1/16	1	$78,608$	20	511
	3	1/32	8	$71,874$	20	472
	3	1/64	64	$68,656$	20	495
	3	1/128	512	$67,084$	20	534
	6	1/16	8	$30,092$	41	$1,493$
	6	1/32	64	$27,514$	43	$1,607$
	6	1/64	512	$26,282$	45	$1,743$
	6	1/128	512	$205,444$	47	$10,191$
2	1	1/32	1	$143,748$	10	313
	1	1/64	8	$137,313$	10	294
	1	1/128	64	$134,168$	10	296
	1	1/256	512	$132,614$	10	302
	3	1/16	1	$78,608$	14	359
	3	1/32	8	$71,874$	15	356
	3	1/64	64	$68,656$	17	413
	3	1/128	512	$67,084$	19	467
	6	1/16	8	$30,092$	19	698
	6	1/32	64	$27,514$	24	903
	6	1/64	512	$26,282$	29	$1,142$
	6	1/128	512	$205,444$	30	$6,320$

TABLE 6.1
V-cycle results- region 1: $\sigma_t = .1$ and $\sigma_a = .05$, region 2: $\sigma_t = 10$ and $\sigma_a = 5$.

From Table 6.1, we see that the number of iterations for the region 1 case is about constant as h is refined. Thus, in region 1, the multigrid algorithm appears to be scaling well spatially. However, the number of iterations increases linearly as the number of spherical harmonic terms is increased (e.g., 9 iterations for $N = 1$ but 20 iterations for $N = 3$). This growth should not be surprising since the size of the system PDE grows quadratically in N. This convergence growth also reveals several properties of the moment coupling. First, the slower but spatially scaled convergence rate for $N = 3$ indicates that relaxation is not effective on some high frequencies of the system PDE. (If the slowly converging components were smooth, then the convergence rate would not scale with h.) Since nodal Gauss-Seidel takes account of the full coupling at a node, these moment-coupling high frequencies must "spread out" spatially. A block Gauss-Seidel that involves the moments over more spatial nodes may give better scaling.

Note that a smooth frequency coupling may also be creeping in as the number of spherical harmonics is increased, as indicated by the slight increase in iterations as h is refined for $N = 6$. This smooth frequency coupling is also more pronounced in region 2 problems, as indicated by the logarithmic growth in the number of iterations as h is refined.

Further features of the moment coupling can be deduced from the preconditioned conjugate gradient results. The overall increase in iteration counts reflects the "strength"

ETNA
Kent State University
etna@mcs.kent.edu

Region	N	h	# Procs	# Unks/Proc	Iter
1	1	1/32	1	$143,748$	17
	1	1/64	8	$137,313$	18
	1	1/128	64	$134,168$	18
	1	1/256	512	$132,614$	18
	3	1/16	1	$78,608$	31
	3	1/32	8	$71,874$	32
	3	1/64	64	$68,656$	33
	3	1/128	512	$67,084$	34
	6	1/16	8	$30,092$	59
	6	1/32	64	$27,514$	60
	6	1/64	512	$26,282$	63
	6	1/128	512	$205,444$	65
2	1	1/32	1	$143,748$	42
	1	1/64	8	$137,313$	54
	1	1/128	64	$134,168$	60
	1	1/256	512	$132,614$	62
	3	1/16	1	$78,608$	19
	3	1/32	8	$71,874$	23
	3	1/64	64	$68,656$	26
	3	1/128	512	$67,084$	29
	6	1/16	8	$30,092$	29
	6	1/32	64	$27,514$	41
	6	1/64	512	$26,282$	52
	6	1/128	512	$205,444$	53

TABLE 6.2
Pcg with 1 $V(1,1)$ preconditioning- region 1: $\sigma_t = .1$ and $\sigma_a = .05$, region 2: $\sigma_t = 10$ and $\sigma_a = 5$.

of the inter-moment coupling since these couplings are not handled well with a block diagonal preconditioner that considers only the intra-moment couplings. However, the peculiar behaviour for $N = 1$ in region 2 is difficult to explain. For $N = 1$, the moments couple only through the boundary functional. Thus, this boundary coupling may be stronger than expected.

Processor scalability is best observed in region 1 results. As the ratio of unknowns per processor is kept roughly constant, the time taken to perform the same amount of computation should be roughly constant. This can be observed in region 1 for $N = 1, 3$, where the number of iterations remains constant as h is refined- e.g., for $N = 3$, with the number of unknowns/processor roughly 70,000, the overall run time is roughly 500 seconds for each refinement.

ii) Region 3. The above multigrid scheme performs poorly; the asymptotic convergence rate approaches 1. Thus, to handle the problematic divergence-free error components, a multiplicative Schwarz smoother is used on the 1-1 moments. The block solves in this smoother simultaneously update the 1-1 moments at the 27 nodes comprising an eight element agglomerate. Table 6.3 tabulates some results. Since the asymptotic convergence rate increases from .58 for N=1 to .68 for N=3, the convergence rate again depends on the number of spherical harmonic terms.

N	h	Rate
1	1/16	0.36
1	1/32	0.56
1	1/64	0.56
3	1/16	0.48
3	1/32	0.64
3	1/64	0.68

TABLE 6.3
$V(1,1)$ convergence rates for region 3: $\sigma_t = 10$ and $\sigma_a = 0.001$.

To investigate locking, we consider the boundary value problem

$$(6.1)\quad \begin{cases} [\Omega\cdot\nabla + \sigma_t I - \sigma_s P]\psi(\mathbf{x},\Omega) = & q & (\mathbf{x},\Omega)\in[0,1]^3\times S^2 \\ \psi(\mathbf{x},\Omega) = & 0 & \mathbf{x}\in\partial[0,1]^3,\ \mathbf{n}\cdot\Omega<0 \end{cases}$$

with exact solution

$$\phi_{lm} = \begin{cases} \sin(2\pi x)\sin(2\pi x)\sin(2\pi z) & l < N = 2 \\ 0 & l = 2, \end{cases}$$

and $\sigma_a = 0.005$. Relative V^1 and H^1 norms of the error in ϕ^h and relative H^1 norm of the error in $(\phi^h_{1,-1}, \phi^h_{1,0}, \phi^h_{1,1})^t$ are given in Table 6.4. The theory in [17] guarantees only order h convergence in the V^1, as confirmed by the results in Table 6.4- as h is halved, the error in the V^1 norm is halved also. Note that irrespective of the magnitude of σ_t, as h is halved, the discretization error in $(\phi^h_{1,-1}, \phi^h_{1,0}, \phi^h_{1,1})^t$ in the H^1 norm decreases about 60%. Since this accuracy improvement is independent of σ_t, we see locking-free error.

σ_t	h	V^1	H^1	H^1:1-1
10	1/16	1.11e-1	1.52e-1	1.61e-1
	1/32	5.49e-2	6.33e-2	6.51e-2
	1/64	2.74e-2	2.93e-2	2.96e-2
50	1/16	1.09e-1	1.77e-1	1.92e-1
	1/32	5.45e-2	7.05e-2	7.43e-2
	1/64	2.73e-2	3.15e-2	3.24e-2
100	1/16	1.09e-1	1.82e-1	1.98e-1
	1/32	5.44e-2	7.43e-2	7.90e-2
	1/64	2.73e-2	3.39e-2	3.55e-2

TABLE 6.4
Locking-free discretization error for region 3 problems: $\sigma_t = 10, 50, 100$, and $\sigma_a = 0.005$. V^1 is the V^1 norm of the total error; H^1 is the H^1 norm of the total error, and H^1:1-1 is the H^1 norm of the error in $(\phi^h_{1,-1}, \phi^h_{1,0}, \phi^h_{1,1})^t$.

iii) Multi-group, Anisotropic Scattering. The last experiment examines the discretization error for a two-group, anisotropic scattering problem:

$$\begin{cases} [\Omega\cdot\nabla + \sigma^g_t I]\,\psi^g & = \sum_{g=1}^{2}\left[\sum_{j=0}^{2}\sigma^{gg'}_{sj}\ P_j\psi^{g'}\right] + q^g & \text{in } [0,1]^3\times S^2 \\ \psi^g & = 0 & \mathbf{n}\cdot\Omega<0,\ x\in\partial[0,1]^3 \end{cases}$$

ETNA
Kent State University
etna@mcs.kent.edu

$$\sigma_{s0}^{gg'} = \begin{pmatrix} 8 & 0 \\ 8 & 8 \end{pmatrix}, \quad \sigma_{s1}^{gg'} = \begin{pmatrix} 2 & 0 \\ 2 & 2 \end{pmatrix}, \quad \sigma_{s2}^{gg'} = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix},$$

$$\sigma_t^1 = \sigma_t^2 = 10.0,$$

and

$$\text{exact solution}: \ \phi_{lm}^g = \begin{cases} \sin(2\pi x)\sin(2\pi y)\sin(2\pi z) & l < N \\ 0 & l = N \end{cases}.$$

This is a down-scattering problem. Table 6.5 shows the multigrid convergence rates for each energy group. Again, these rates reflect spatial scaling, and growth with respect to the number of spherical harmonic terms. Also, from Table 6.6, again as h is halved, the V^1 norm of the error is also halved. Thus, we see order h discretization accuracy.

Group	N	h	Iter
1	3	1/32	13
	3	1/64	13
	3	1/128	13
	6	1/32	16
	6	1/64	18
	6	1/128	18
2	3	1/32	13
	3	1/64	13
	3	1/128	13
	6	1/32	16
	6	1/64	18
	6	1/128	18

TABLE 6.5
$V(1,1)$ convergence rate for multigroup, anisotropic scattering problem.

Group	h	$N=3$	$N=6$
1	1/32	3.88e-2	3.37e-2
	1/64	1.89e-2	1.63e-2
	1/128	9.39e-3	8.08e-3
2	1/32	3.88e-2	3.38e-2
	1/64	1.89e-2	1.63e-2
	1/128	9.39e-3	8.08e-3

TABLE 6.6
Relative V^1 norm of error for multigroup, anisotropic scattering problem.

7. Conclusion. In this paper, we presented two system multigrid algorithms for solving the anisotropic scattering Boltzmann equation formulated as a FOSLS minimization problem. Used in the inner loop of a block-group Gauss-Seidel iteration, either algorithms gives an efficient solver for many important multi-group, anisotropic scattering Boltzmann equations. Morever, using a multiplicative Schwarz smoother instead of a nodal Gauss-Seidel smoother, we are able to get an efficient algorithm for the region 3 1-1 block. Because of the favorable scaling of this time-depedent incompressible elasticity-like problem, locking seems to be less severe than as in the case of general incompressible elasticity problems. Numerical results demonstrate that locking does not occur even for low-order, conforming finite elements. Other numerical results demonstrate that these new multigrid algorithms scale better than the PCG iteration examined in [9] for region 1 and 2 problems. However, numerical results also indicate that the higher moments couple through high frequency modes. To handle these error modes better, more sophisticated smoothers need to be designed. This is the topic of current research.

REFERENCES

[1] R. A. ADAMS, *Sobolev Spaces*, Academic Press, 1975.
[2] R. E. ALCOUFFE, A. BRANDT, J. E. DENDY, AND J. W. PAINTER, *The multi-grid method for the diffusion equation with strongly discontinuous coefficients*, SIAM J. Sci. Stat. Comput., 2 (1981), pp.430-454.
[3] D. N. ARNOLD, R. S. FALK, AND R. WINTHER, *Preconditioning in $H(div)$ and applications*, Math. Comp., 66 (1997), pp. 957-984.
[4] I. BABUSKA AND M. SURI, *Locking effects in the finite element approximation of elasticity problems*, Numer. Math., 62 (1992), pp.439-463.
[5] D. BRAESS, *Finite elements theory, fast solvers, and applications in solid mechanics*, Cambridge University Press, Cambridge, 1997.
[6] S. C. BRENNER AND L. R. SCOTT, *The Mathematical Theory of Finite Element Methods*, Springer-Verlag, New York, 1994.
[7] F. BREZZI AND M. FORTIN, *Mixed and Hybrid Finite Element Methods*, Springer-Verlag, New York, 1991.
[8] Z. CAI, T. MANTEUFFEL, AND S. McCORMICK, *First-order system least squares for partial differential equations: Part II*, SIAM J. Numer. Anal., 34 (1997), pp. 425-454.
[9] B. CHANG AND B. LEE, *Space-Angle First-Order System Least-Squares (FOSLS) for the Linear Boltzmann Equation*, preprint.
[10] J. E. DENDY, *Black box multigrid*, J. Comp. Phys., 48 (1982), pp.366-386.
[11] R. HIPTMAIR, *Multigrid method for $H(div)$ in three dimensions*, ETNA, 6 (1997), pp.133-152.
[12] K. KOBAYASHI, N. SUGIMURA, AND Y. NAGAYA, *3D radiation transport benchmarks for simple geometries with void regions*, preprint.
[13] E. W. LARSEN, J. E. MOREL, AND W. F. MILLER, *Asymptotic solutions of numerical transport problems in optically thick, diffusive regimes*, J. Comp. Phys., 69 (1987), pp.283-324.
[14] E. E. LEWIS AND W. F. MILLER, *Computational Methods of Neutron Transport*, John Wiley, New York, 1984.
[15] R. L. LIBOFF, *Introductory Quantum Mechanics*, Holden-Day, Inc., San Francisco, 1980.
[16] T. A. MANTEUFFEL AND K. J. RESSEL, *Least-squares finite-element solution of the neutron transport equation in diffusive regimes*, SIAM J. Numer. Anal., 35 (1998), pp. 806-835.
[17] T. A. MANTEUFFEL, K. J. RESSEL, AND G. STARKE, *A boundary functional for the least-squares finite element solution of neutron transport problems*, SIAM J. Numer. Anal., 37 (2000), pp.556-586.
[18] T. A. MANTEUFFEL, private communication.
[19] J. D. MOULTON, J. E. DENDY, JR., AND J. M. HYMAN, *The black box multigrid numerical homogenization algorithm*, J. Comp. Phys., 142 (1998), pp.80-108.
[20] J. J. SAKURAI, *Modern Quantum Mechanics*, Addison-Wesley, 1994.
[21] J. SCHOBERL, *Multigrid methods for a parameter dependent problem in primal variables*, Numer. Math., 84 (1999), pp.97-119.
[22] P. S. VASSILEVSKI AND J. WANG, *Multilevel iterative methods for mixed finite element discretizations of elliptic problems*, Numer. Math., 63 (1992), pp. 503-520.
[23] P. M. D. ZEEUW, *Matrix-dependent prolongations and restrictions in a blackbox multigrid solver*, J. Comp. Appl. Math., 33 (1990), pp.1-27.