A Fast Multigrid Algorithm for Isotropic Transport Problems I: Pure Scattering *

T. Manteuffel[†], S. McCormick[†], J. Morel[‡], S. Oliveira[§] and G. Yang[§]

Abstract

A multigrid method for solving the 1-D slab-geometry S_N equations with isotropic scattering and no absorption is presented. This scheme is highly compatible with massively parallel computer architectures and represents a first step towards similar multigrid methods for the S_N equations in curvilinear and multi-dimensional geometries. Extensive theoretical analyses are given for our scheme which indicate that it is extremely efficient. In fact, the method is so efficient that it very nearly represents an exact solution technique. Results from calculations are presented which validate the theoretical results. The case with absorption is treated in a sequel to this paper [12].

§1. Introduction

In this paper we describe a fast method for solving the equations used to model the transport of neutral particles with isotropic scattering in slab geometry. These problems are important in many application such as nuclear reactor design, radiation therapy in medical science, radiation effects on global weather, and are a fundamental part of many algorithms used to model more complicated applications such as satellite electronics shielding. For example, a multigrid technique for the solution of highly forward-peaked anisotropic scattering in slab geometry requires the solution of an isotropic problem at the coarsest level (Morel and Manteuffel [15]). The main focus of this paper is the extensive theoretical results which we have obtained for our multigrid method. The matrix analysis techniques which we use are quite different from the Fourier analysis techniques traditionally used in numerical transport theory (e.g., see Morel and Manteuffel [15]). Since these matrix techniques yield information which is not usually available from a Fourier analysis, they represent a valuable alternative to Fourier analysis.

*This work was supported by AFOSR under grant AFOSR 86-0126, the NSF under grant DMS-8704169,the DOE under grant DE-FG02-90ER25086 and Los Alamos National Laboratory.

[†]Program in Applied mathematics, University of Colorado at Boulder, Boulder, CO

[‡]Computer Research Group (C-3), Los Alamos National Laboratory, Los Alamos, NM

[§]Center for Computational Mathematics, University of Colorado of Denver, Denver, CO

The linear Boltzman equation, which is used to model neutral-particle transport, reduces to the following equation in slab geometry with isotropic scattering. The physical domain is a semi-infinite slab of width $b - a$ in the x dimension. Although three dimensional, we assume the flux of particles is independent of the y and z coordinates. Thus, the Linear Boltzman equation reduces to

$$\mu \frac{\partial \psi}{\partial x} + \sigma_t \psi = \frac{\sigma_s}{2} \int_{-1}^{1} \psi(x, \mu') d\mu' + q(x, \mu), \tag{1.1}$$

for $x \epsilon (a, b)$, $\mu \epsilon [-1, 1]$, where $\psi(x, \mu)$ represents the flux of particles at position x traveling at an angle θ from the x-axis ($\mu = cos(\theta)$). Here $\sigma_t \, dx$ represents the expected number of interactions (absorptive or scattering) that a particle will have in traveling a distance dx. In a similar manner, $\sigma_s \, dx$ represents the expected number of scattering interactions while $\sigma_a = \sigma_t - \sigma_s$ represents the expected number of absorptive interactions. Finally, $q(x, \mu)$ represents the particle source.

The solution is well defined if the flux of particles entering the slab is given as boundary conditions:

$$\psi(a, \mu) = g_a(\mu), \quad \psi(b, -\mu) = g_b(\mu) \quad \mu \epsilon (0, 1). \tag{1.2}$$

The problem (1.1) becomes difficult to solve in the optically dense or thick diffusion limit. Physically, the thick limit implies that the mean-free-path between collisions that is small compared to the width of the slab. If, in addition, the material allows very little absorption, the model resembles diffusion away from boundaries and sources. Mathematically, we have

$$\sigma_t \to \infty, \quad \frac{\sigma_s}{\sigma_t} \to 1. \tag{1.3}$$

Dividing (1.1) by σ_t and taking the limits in (1.3) yields

$$\psi(x, \mu) = \frac{1}{2} \int_{-1}^{1} \psi(x, \mu') d\mu', \tag{1.4}$$

which admits any $\psi(x, \mu)$ that is independent of μ. Thus, in this limit (1.1) is singularly perturbed with a large near null space. The problems near the thick diffusion limit are difficult to solve for traditional transport methods.

Numerical approximations of (1.1) suffer several difficulties in this limit. Many difference schemes yield inaccurate solutions in this limit even for well behaved solutions. For a discussion of this issue see Larsen and Morel[8]. One difference scheme that behaves well in the thick limit is the Modified Linear Discontinuous Scheme(MLD)([8]). Not only does it give the proper behavior in the thick limit but it is very accurate. The algorithm described in this paper will address the solution of the transport equation with S_N discretization in angle (See Section 2) using MLD in space.

Even when the difference scheme behaves properly in the thick limit it may be difficult to solve the resulting discrete equations. The Diffusion Synthetic Acceleration(DSA)

scheme was developed to address this difficulty(Alcouffe[2]). DSA is based upon the recognition that in the thick limit the transport of particles becomes diffusive in nature. The flux is nearly independent of direction except near the boundarys and sources. DSA involves solving a diffusion equation as a preconditioning for the transport equations (See Faber and Manteuffel[5]). In theory, it achieves a convergence factor bounded by approximately .23 per iteration independent of σ_t or the cell width h.

In this paper we present a multigrid algorithm that is much faster than DSA and avoids some of the problems associated with DSA. In general, the DSA algorithm requires a certain degree of consistency between the S_N spatial discretization scheme and the diffusion spatial discretization scheme. In recent work Larsen [6], Morel and Larsen [14] and Adams and Martin [1] have derived methods for finding consistent diffusion schemes. In [1] a scheme was developed in which the convergence factor goes to zero as $\sigma_t h$ gets large. However, DSA works because the transport model becomes diffusive in the thick limit. Such problems are usually amenable to multigrid techniques directly. Moreover, the difficulties associated with deriving and solving consistent DSA diffusion equations increase with higher dimensions.

The multigrid techniques described here have a natural generalization to higher dimensions and, as mentioned previously, can be used in conjunction with the multigrid algorithm for anisotropic problems described in Morel and Manteuffel[15]. While any generalization to higher dimensions is fraught with peril, we feel that the multigrid algorithm algorithms presented here can be successfully extended.

Although similar in spirit, our multigrid method is much more efficient than those of Nowak [16] [17], Nowak and Larsen [18], Barnett, Morel, and Harris [3] and Nowak, Morel and Harris [19]. These methods require an expensive relaxation and do not achieve a convergance factor that approaches zero in the thick diffusion limit.

In this paper, we concentrate on pure scattering, that is, no absorption, which is modeled by setting $\sigma_s = \sigma_t$. Convergence proofs for both the thick and thin limits are developed. The efficacy of the algorithm is demonstrated with numerical examples. The model with absorption, i. e., $\sigma_s < \sigma_t$, is addressed in a sequel to this paper, [12]. In this paper piecewise linear discontinuous elements are used for the spatial discretization. In [12] piecewise kinked-linear discontinuous elements are used in the thick cell limit. The proofs in this paper are used to develop asymptotic estimates for the algorithm in [12].

The remainder of this paper is organized as follows. In Section 2 we describe the discrete transport equation. We use an S_N approximation in direction (angle), and modified piecewise linear discontinuous elements (MLD) for the spatial discretization. In Section 3 we develop a two-cell, red/black, block μ-line relaxation. The fact that the scattering operator is rank-1 allows for a very efficient μ-line relaxation. Letting h_i be the width of cell i, we show

a) for $\max(\sigma_t h_{2i-1}, \sigma_t h_{2i}) \ll 1$ at two neighboring cells $2i-1$ and $2i$, the μ-line relaxation leaves the error linear across the two cells up to order $O(\max(\sigma_t^2 h_{2i-1}^2, \sigma_t^2 h_{2i}^2))$.

b) for $\min(\sigma_t h_{2i-1}, \sigma_t h_{2i}) \gg 1$, the μ-line relaxation leaves the error linear across the two cells up to order $O(\max(\frac{1}{\sigma_t h_{2i-1}}, \frac{1}{\sigma_t h_{2i}}))$.

In Section 4 we combine the relaxation with a multigrid algorithm and derive two-grid convergence rates, ρ;

a) for $\max\limits_i\ (\sigma_t h_i) \ll 1, \quad \rho = O(\max\limits_i\ (\sigma_t h_i)^2)$,

b) for $\min\limits_i\ (\sigma_t h_i) \gg 1, \quad \rho = O(\max\limits_i\ (\frac{1}{\sigma_t h_i}))$.

In Section 5.1 we present numerical results for our multigrid method for both uniform and highly nonuniform grids. They yield the following convergence factor,ρ, for a V(1,1) cycle:

a) for $\max\limits_i\ (\sigma_t h_i) \ll 1, \quad \rho = O(\max\limits_i\ (\sigma_t h_i)^3)$,
b) for $\min\limits_i\ (\sigma_t h_i) \gg 1, \quad \rho = O(\frac{1}{\min\limits_i\ (\sigma_t h_i)^2})$,

with a slowest rate of $\rho \sim 0.01$ occurring for $\sigma_t h \sim 0.01$. These rates make a single $V(1,1)$ cycle nearly an exact solver, especially in the thick and thin limits. The cost of a $V(1,1)$ cycle is on the same order as 4 sweeps across the computational grid. This represents a significant savings over DSA. Moreover, the algorithm was designed to be highly parallel. We have implemented the algorithm on the Thinking Machines Inc., CM-2 at Los Alamos National Labs. These results appear in (Manteuffel et.al, [10], [11]). In Section 5.2 we present numerical results for a DSA scheme for the MLD equations. Section 6 contains concluding remarks.

§2. The Discrete Model

The angle dependence in (1.1) is discretized by a S_N approximation(c.f. Lewis and Miller [9]), that is, we assume

$$\psi(x,\mu) = \sum_{l=0}^{N-1} (2l+1)\phi_l(x)p_l(\mu), \tag{2.1}$$

where $p_l(\mu)$ is the l^{th} Legendre polynomial and $(2l+1)$ is a normalization factor. If standard Galerkin formulation is used, a system of equations, for the $\phi_l(x)'s$, called the moment equations, is established. The moments may be found by

$$\phi(x) = \frac{1}{2}\int_{-1}^{1} \psi_l(x,\mu')p_l(\mu')d\mu' = \sum_{k=1}^{N} \omega_k \psi_l(x,\mu_k)p_l(\mu_k)d\mu', \tag{2.2}$$

where $\omega_1, ..., \omega_N$ and $\mu_1, ..., \mu_N$ are the Gauss quadrature weights and points respectively. The second equality is exact because of the assumption (2.1). Thus, finding $\psi_j(x) = \psi(x,\mu_j)$ for j=1,...,N is equivalent to finding the moments. A simple transformation of the moment equations yields a system for the $\psi_j(x)$'s, called the flux equations, as follows

$$\mu_j \frac{\partial \psi_j(x)}{\partial x} + \sigma_t \psi_j(x) = \sigma_s \sum_{k=1}^{N} \omega_k \psi_k(x) + q_j(x), \tag{2.3}$$

for j=1,...,N.

For our purposes here we choose to differentiate the positive and negative direction. Assume N is even and let $n = N/2$, the Gauss quadrature points are symmetric about the origin and can be denoted as $-\mu_n <, ..., < -\mu_1 < 0 < \mu_1 <, ..., \mu_n$. Likewise, the weight associated with $-\mu_j$ will equal the weight associated with μ_j. We define

$$\psi_j^+(x) = \psi(x, \mu_j), \tag{2.4a}$$

$$\psi_j^-(x) = \psi(x, -\mu_j), \tag{2.4b}$$

for $j = 1, ..., n$. Equation(2.3) becomes the pair of equations

$$\mu_j \frac{\partial \psi_j^+(x)}{\partial x} + \sigma_t \psi_j^+(x) = \sigma_s \sum_{k=1}^{n} \omega_k(\psi_k^+(x) + \psi_k^-(x)) + q_j^+(x), \tag{2.5a}$$

$$\mu_j \frac{\partial \psi_j^-(x)}{\partial x} + \sigma_t \psi_j^-(x) = \sigma_s \sum_{k=1}^{n} \omega_k(\psi_k^+(x) + \psi_k^-(x)) + q_j^-(x), \tag{2.5b}$$

for $j = 1, ..., n$. The boundary conditions (1.2) become

$$\psi_j^+(a) = g_a(\mu_j), \tag{2.6a}$$

$$\psi_j^-(b) = g_b(\mu_j), \tag{2.6b}$$

for $j = 1, ..., n$.

The space dependence is now discretized by the Modified Linear Discontinuous difference scheme (Larsen and Morel[8]). The MLD scheme can be derived by using finite element methods. We first describe the Linear Discontinuous difference scheme (LD)[8]. Consider a set of grid points given by

$$a = x_{\frac{1}{2}} < x_{\frac{3}{2}} < ... < x_{m+\frac{1}{2}} = b. \tag{2.7}$$

These represent cell edges. The centers are given by

$$x_i = \frac{1}{2}(x_{i-\frac{1}{2}} + x_{i+\frac{1}{2}}) \quad \text{for} \quad i = 1, ..., m, \tag{2.8}$$

and the cell width is $h_i = x_{i+\frac{1}{2}} - x_{i-\frac{1}{2}}$ and m is the number of cells. We start by forming piecewise linear trial spaces. However, a different space is used for $\psi_j^+(x)$ than for $\psi_j^-(x)$. For $\psi_j^+(x)$(for each j), we choose break points at

$$x_{\frac{1}{2}} < \hat{x}_1^+ < x_{\frac{3}{2}} < \ldots < x_{i-\frac{1}{2}} < \hat{x}_i^+ < x_{i+\frac{1}{2}} < \ldots \tag{2.9a}$$

with $\hat{x}_i^+ = x_{i-\frac{1}{2}} + \epsilon_i$, while for $\psi_j^-(x)$(for each j) we choose break points at

$$x_{\frac{1}{2}} < \hat{x}_1^- < x_{\frac{3}{2}} < \ldots < x_{i-\frac{1}{2}} < \hat{x}_i^- < x_{i+\frac{1}{2}} < \ldots \tag{2.9b}$$

with $\hat{x}_i^- = x_{i+\frac{1}{2}} - \epsilon_i$. Here ϵ_i is small. Bases elements for the two spaces, which we denote as $V_{h,\epsilon}^+$ and $V_{h,\epsilon}^-$, are depicted in Figures 1 and 2 respectively.

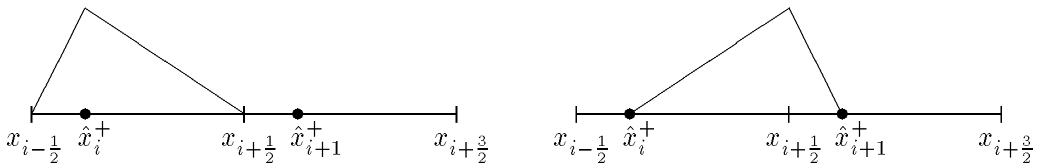

Figure 1

Bases for $V_{h,\epsilon}^+$

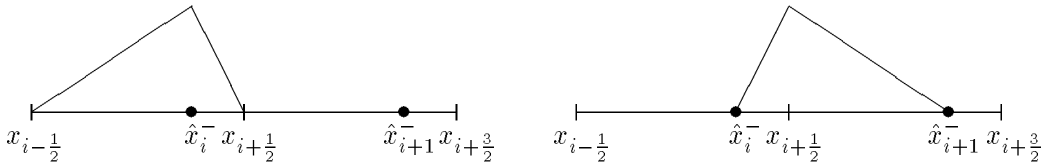

Figure 2

Bases for $V_{h,\epsilon}^-$

We assume that $\psi_j^+(x)$ and $\psi_j^-(x)$ can be expressed as a linear combination of the bases elements in $V_{h,\epsilon}^+$ and $V_{h,\epsilon}^-$ respectively. We denote

$$\psi_{i+\frac{1}{2},j}^+ = \psi_j^+(x_{i+\frac{1}{2}}), \quad \psi_{i+\frac{1}{2},j}^- = \psi_j^-(x_{i+\frac{1}{2}}), \tag{2.10a}$$

$$\hat{\psi}_{ij}^+ = \psi_j^+(\hat{x}_i^+), \quad \hat{\psi}_{ij}^- = \psi_j^-(\hat{x}_i^-), \tag{2.10b}$$

We also denote

$$\psi_{i,j}^+ = \psi_j^+(x_i), \quad \psi_{i,j}^- = \psi_j^-(x_i), \tag{2.10c}$$

It is important to note that $\hat{\psi}^+_{i,j}$ and $\hat{\psi}^-_{i,j}$ can be expressed in terms of values of $\psi^+_j(x)$ and $\psi^-_j(x)$ at the cell edges and cell center. Since $\psi^+_j(x)$ and $\psi^-_j(x)$ are assumed to be linear, $\hat{\psi}^+_{i,j}$ and $\hat{\psi}^-_{i,j}$ can be found as an extrapolation from the center and edge. We have

$$\hat{\psi}^+_{i,j} = 2(\frac{h_i - \epsilon_i}{h_i})\psi^+_{i,j} - (\frac{h_i - 2\epsilon_i}{h_i})\psi^+_{i+\frac{1}{2},j}, \tag{2.11a}$$

$$\hat{\psi}^-_{i,j} = 2(\frac{h_i - \epsilon_i}{h_i})\psi^-_{i,j} - (\frac{h_i - 2\epsilon_i}{h_i})\psi^-_{i-\frac{1}{2},j}. \tag{2.11b}$$

The Linear Discontinuous (LD) difference Scheme is derived by applying a standard Galerkin Formulation, then taking the limit as $\max_i (\epsilon_i) \to 0$. One way to construct the equations is to plug the trial space representation into $(2.5a)$ and $(2.5b)$ and integrate over each cell against the constant function and then against the linear function, $(x - x_i)$. This will yield a set of equations involving ϵ_i. The final equations are found by setting ϵ_i=0.

The Modified Linear Discontinuous Scheme(MLD) is found in a similar manner. As in the LD scheme, $(2.5a)$ and $(2.5b)$ are integrated over each cell against the constant function. Then after setting $\epsilon_i = 0$ we get the following equations

(2.12a)

$$\mu_j(\psi^+_{i+\frac{1}{2},j} - \psi^+_{i-\frac{1}{2}j}) + \sigma_t(\frac{\hat{\psi}^+_{ij} + \psi^+_{i+\frac{1}{2}j}}{2})h_i = \sigma_s \sum_{k=1}^{n} \omega_k(\hat{\psi}^+_{i,k} + \psi^+_{i+\frac{1}{2}k} + \hat{\psi}^-_{i,k} + \psi^-_{i-\frac{1}{2}k})h_i + q^+_{ij}h_i,$$

(2.12b)

$$\mu_j(\psi^-_{i+\frac{1}{2},j} - \psi^-_{i-\frac{1}{2}j}) + \sigma_t(\frac{\hat{\psi}^-_{ij} + \psi^-_{i-\frac{1}{2}j}}{2})h_i = \sigma_s \sum_{k=1}^{n} \omega_k(\hat{\psi}^+_{i,k} + \psi^+_{i+\frac{1}{2}k} + \hat{\psi}^-_{i,k} + \psi^-_{i-\frac{1}{2}k})h_i + q^-_{ij}h_i.$$

These are referred to as the balance equations. Instead of integration against linear functions, as in LD, now the equations $(2.5a)$ are collocated at $\hat{x}^-_i$, equations $(2.5b)$ are collocated at $\hat{x}^+_i$, and the limit is taken as $\max_i \epsilon_i \to 0$. This yields the following equations referred to as edge equations.

$$\mu_j(\frac{\psi^+_{i+\frac{1}{2},j} - \hat{\psi}^+_{ij}}{h_i}) + \sigma_t\psi^+_{i+\frac{1}{2}j} = \sigma_s \sum_{k=1}^{n} \omega_k(\psi^+_{i+\frac{1}{2},k} + \hat{\psi}^-_{ik}) + q^+_{i+\frac{1}{2}j}, \tag{2.13a}$$

$$\mu_j(\frac{\hat{\psi}^-_{ij} - \psi^-_{i-\frac{1}{2}j}}{h_i}) + \sigma_t\psi^-_{i-\frac{1}{2}j} = \sigma_s \sum_{k=1}^{n} \omega_k(\hat{\psi}^+_{ik} + \psi^-_{i-\frac{1}{2}k}) + q^-_{i-\frac{1}{2}j}. \tag{2.13b}$$

This may be formalized as integration of $(2.5a)$ against $\delta(x - \hat{x}^-_i)$ and $(2.5b)$ against $\delta(x - \hat{x}^+_i)$. Note from $(2.11a, b)$ that we may replace $\hat{\psi}^+_{ik}$ and $\hat{\psi}^-_{ik}$ with $2\psi^+_{ik} - \psi^+_{i+\frac{1}{2}}$ and

$2\psi_{ik}^- - \psi_{i-\frac{1}{2}}^-$ respectively. The boundary conditions are

$$\psi_{\frac{1}{2}j}^+ = g_{aj}^+, \quad \psi_{m+\frac{1}{2}j}^- = g_{bj}^-, \ j = 1, 2, ..., n. \tag{2.14}$$

We see from the edge equations and balance equations that it is the product of σ_t and h_i that will determine the property of the discretized equations. We will use $\sigma_t h_i$ as a parameter in our discussion and set $\gamma = \frac{\sigma_s}{\sigma_t} = 1$.

In block matrix form, these equations become

$$B_i(\underline{\psi}_{i+\frac{1}{2}}^+ - \underline{\psi}_{i-\frac{1}{2}}^+) + \underline{\psi}_i^+ = R(\underline{\psi}_i^+ + \underline{\psi}_i^-) + \underline{q}_i^+, \tag{2.15a}$$

$$2B_i(\underline{\psi}_{i+\frac{1}{2}}^+ - \underline{\psi}_i^+) + \underline{\psi}_{i+\frac{1}{2}}^+ = R(\underline{\psi}_{i+\frac{1}{2}}^+ + 2\underline{\psi}_i^- - \underline{\psi}_{i-\frac{1}{2}}^-) + \underline{q}_{i+\frac{1}{2}}^+, \tag{2.15b}$$

for i = 1,...,m, and

$$B_i(\underline{\psi}_{i-\frac{1}{2}}^- - \underline{\psi}_{i+\frac{1}{2}}^-) + \underline{\psi}_i^- = R(\underline{\psi}_i^+ + \underline{\psi}_i^-) + \underline{q}_i^-, \tag{2.15c}$$

$$2B_i(\underline{\psi}_{i-\frac{1}{2}}^- - \underline{\psi}_i^-) + \underline{\psi}_{i-\frac{1}{2}}^- = R(\underline{\psi}_{i-\frac{1}{2}}^- + 2\underline{\psi}_i^+ - \underline{\psi}_{i+\frac{1}{2}}^+) + \underline{q}_{i-\frac{1}{2}}^-, \tag{2.15d}$$

for $i = 1, ..., m$, with boundary conditions

$$\underline{\psi}_{\frac{1}{2}}^+ = \underline{g}_a^+, \quad \underline{\psi}_{m+\frac{1}{2}}^- = \underline{g}_b^-, \tag{2.15e}$$

where B_i and R are defined as

$$B_i = \begin{bmatrix} \frac{\mu_1}{\sigma_t h_i} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \frac{\mu_n}{\sigma_t h_i} \end{bmatrix}, \qquad R = \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix} \begin{bmatrix} \omega_1 & \cdots & \omega_n \end{bmatrix} = \underline{1}\, \underline{\omega}^T, \tag{2.16a}$$

and

$$\underline{\psi}_{i+\frac{1}{2}}^+ = (\psi_{i+\frac{1}{2}1}^+, ..., \psi_{i+\frac{1}{2}n}^+)^T, \quad \underline{\psi}_i^+ = (\psi_{i1}^+, ..., \psi_{in}^+)^T, \tag{2.16b}$$

$$\underline{\psi}_{i+\frac{1}{2}}^- = (\psi_{i+\frac{1}{2}1}^-, ..., \psi_{i+\frac{1}{2}n}^-)^T, \quad \underline{\psi}_i^- = (\psi_{i1}^-, ..., \psi_{in}^-)^T. \tag{2.16c}$$

We can write equation (2.15) in matrix form as

$$\begin{bmatrix} A_1 & -C_1 & & & & \\ -D_2 & A_2 & -C_2 & & & \\ & \ddots & \ddots & \ddots & & \\ & & -D_i & A_i & -C_i & \\ & & & \ddots & \ddots & \ddots \\ & & & & -D_m & A_m \end{bmatrix} \begin{bmatrix} \underline{\Psi}_1 \\ \underline{\Psi}_2 \\ \vdots \\ \underline{\Psi}_i \\ \vdots \\ \underline{\Psi}_m \end{bmatrix} = \begin{bmatrix} \underline{Q} \end{bmatrix}, \tag{2.17}$$

where $\underline{Q}$ is the right-hand side and

$$A_i = \begin{bmatrix} I+2B_i-R & -2R & -2B_i & R \\ 0 & I-R & -R & B_i \\ B_i & -R & I-R & 0 \\ R & -2B_i & -2R & I+2B_i-R \end{bmatrix}, \tag{2.18a}$$

$i = 1, ..., m,$

$$C_i = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ B_i & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, i = 1, ..., m-1, \tag{2.18b}$$

$$D_i = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & B_i \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, i = 2, ..., m, \tag{2.18c}$$

$$\underline{\Psi}_i = (\underline{\psi}^-_{i-\frac{1}{2}}, \underline{\psi}^+_i, \underline{\psi}^-_i, \underline{\psi}^+_{i+\frac{1}{2}})^T, i = 1, ..., m. \tag{2.18d}$$

In the following section, we will examine the x-line relaxation, also know as a transport sweep or source iteration, and the μ-line relaxation and their properties.

§3 Relaxation

Suppose the right-hand sides of (2.15a-d) were known. The positive angles would then satisfy a system of equations involving the matrix

$$\begin{bmatrix} I & & & & & & \\ -B_1 & I & B_1 & & & & \\ & -2B_1 & I+2B_1 & & & & \\ & & -B_2 & I & B_2 & & \\ & & & -2B_2 & I+2B_2 & & \\ & & & & & & \ddots \end{bmatrix} \begin{bmatrix} \underline{\psi}^+_{\frac{1}{2}} \\ \underline{\psi}^+_1 \\ \underline{\psi}^+_{\frac{3}{2}} \\ \underline{\psi}^+_2 \\ \underline{\psi}^+_{\frac{5}{2}} \\ \vdots \end{bmatrix}. \tag{3.1}$$

The first block equation represents the boundary conditions. This block lower triangular matrix could be solved in a forward sweep by inverting the block matrices as follows:

$$\begin{bmatrix} I & B_i \\ -2B_i & I+2B_i \end{bmatrix}^{-1} = (I+2B_i+2B_i^2)^{-1} \begin{bmatrix} I+2B_i & -B_i \\ 2B_i & I \end{bmatrix}. \tag{3.2}$$

Since B_i is diagonal, this is a trivial operation. Likewise, the negative angles could be found in a backward sweep.

This process represents x-line relaxation (also known as a source iteration or transport sweep). If we reorder the unknowns by cells and examine the effect of a complete x-line relaxation, the resulting error equation for a single cell can be written as

$$
\begin{bmatrix} I+2B_i & 0 & -2B_i & 0 \\ 0 & I & 0 & -B_i \\ -B_i & 0 & I & 0 \\ 0 & -2B_i & 0 & I+2B_i \end{bmatrix} \begin{bmatrix} \underline{e}^-_{i-\frac{1}{2}} \\ \underline{e}^+_i \\ \underline{e}^-_i \\ \underline{e}^+_{i+\frac{1}{2}} \end{bmatrix}^{l+1} = \tag{3.3}
$$

$$
\begin{bmatrix} R & 2R & 0 & -R \\ 0 & R & R & 0 \\ 0 & R & R & 0 \\ -R & 0 & 2R & R \end{bmatrix} \begin{bmatrix} \underline{e}^-_{i-\frac{1}{2}} \\ \underline{e}^+_i \\ \underline{e}^-_i \\ \underline{e}^+_{i+\frac{1}{2}} \end{bmatrix}^{l} + \begin{bmatrix} 0 \\ B_i\underline{e}^+_{i-\frac{1}{2}} \\ B_i\underline{e}^-_{i+\frac{1}{2}} \\ 0 \end{bmatrix}^{l+1} .
$$

Here l is the relaxation step. When $\sigma_t h_i \gg 1$, x-line relaxation is not effective. Equation (2.16a) reveals that, when $\sigma_t h_i \to \infty$, B_i approaches zero. Then, an x-line relaxation nearly reproduces the component of the error that is independent of angle. Since this does not affect any spatial frequency, the resulting error is not smooth in space. Any error that is independent of angle will not be suppressed regardless of its spatial frequency. This makes x-line relaxation inappropriate for use in a multigrid scheme, since even if $\sigma_t h_i$ is small on the finest grid, it will become large on coarser grids.

On the other hand, we can rearrange equation (2.17) to reveal the form of two-cell red-black block μ - line relaxations. By two-cell red-black block μ-line relaxation, we mean that at cells i and $i+1$, we set $\underline{\psi}^+_{i-\frac{1}{2}}$ and $\underline{\psi}^-_{i+\frac{3}{2}}$ as boundary values and solve all other interior values in cells i and $i+1$ simultaneously. This relaxation is carried out in a red-black ordering. Since the relaxation uses a red-black ordering, we can look at each 2-cell pair individually. For 2-cell pair i and $i+1$, the errors at these two cells after the relaxation will be

$$
\begin{bmatrix} \underline{e}^-_{i-\frac{1}{2}} \\ \underline{e}^+_i \\ \underline{e}^-_i \\ \underline{e}^+_{i+\frac{1}{2}} \\ \underline{e}^-_{i+\frac{1}{2}} \\ \underline{e}^+_{i+1} \\ \underline{e}^-_{i+1} \\ \underline{e}^+_{i+\frac{3}{2}} \end{bmatrix} = \begin{bmatrix} A_i & -C_i \\ -D_{i+1} & A_{i+1} \end{bmatrix}^{-1} \begin{bmatrix} 0 \\ B_i\underline{e}^+_{i-\frac{1}{2}} \\ 0 \\ 0 \\ 0 \\ 0 \\ B_{i+1}\underline{e}^-_{i+\frac{3}{2}} \\ 0 \end{bmatrix} . \tag{3.4}
$$

The inversion of the $\begin{bmatrix} A_i & -C_i \\ -D_{i+1} & A_{i+1} \end{bmatrix}$ is inexpensive since R is rank one and I, B_i, and

B_{i+1} are diagonal matrices. To show this, we write matrix $\begin{bmatrix} A_i & -C_i \\ -D_{i+1} & A_{i+1} \end{bmatrix}$ as

(3.5)

$$\begin{bmatrix} A_i & -C_i \\ -D_{i+1} & A_{i+1} \end{bmatrix} = \begin{bmatrix} I+2B_i & 0 & -2B_i & 0 & & & & \\ 0 & I & 0 & B_i & & & & \\ B_i & 0 & I & 0 & -B_i & & & \\ 0 & -2B_i & 0 & I+2B_i & & & & \\ & & & & I+2B_{i+1} & 0 & -2B_{i+1} & 0 \\ & & & -B_{i+1} & 0 & I & 0 & B_{i+1} \\ & & & & B_{i+1} & 0 & I & 0 \\ & & & & 0 & -2B_{i+1} & 0 & I+2B_{i+1} \end{bmatrix}$$

$$- \begin{bmatrix} 2\underline{1} & 0 & & \\ \underline{1} & \underline{1} & & \\ \underline{1} & \underline{1} & & \\ 0 & 2\underline{1} & & \\ & & 2\underline{1} & 0 \\ & & \underline{1} & \underline{1} \\ & & \underline{1} & \underline{1} \\ & & 0 & 2\underline{1} \end{bmatrix} \begin{bmatrix} \frac{1}{2}\underline{w}^T & \underline{w}^T & 0 & -\frac{1}{2}\underline{w}^T & & & & \\ -\frac{1}{2}\underline{w}^T & 0 & \underline{w}^T & \frac{1}{2}\underline{w}^T & & & & \\ & & & & \frac{1}{2}\underline{w}^T & \underline{w}^T & 0 & -\frac{1}{2}\underline{w}^T \\ & & & & -\frac{1}{2}\underline{w}^T & 0 & \underline{w}^T & \frac{1}{2}\underline{w}^T \end{bmatrix},$$

where $\underline{1} = (1, 1, ..., 1)^T$ and $\underline{w}^T = (\omega_1, \omega_2, ..., \omega_n)$. The first matrix in (3.5) can be easily inverted since it is sparse and its components I, B_i, and B_{i+1} are diagonal matrices. The second term is a rank four matrix. If we denote (3.5) as

$$\begin{bmatrix} A_i & -C_i \\ -D_{i+1} & A_{i+1} \end{bmatrix} = A_0 - VW^T, \tag{3.6}$$

where A_0 is the first matrix in (3.5) and V, W^T are the two rank four matrices in (3.5), then the Sherman-Morrison formulas yield

$$\begin{bmatrix} A_i & -C_i \\ -D_{i+1} & A_{i+1} \end{bmatrix}^{-1} = A_0^{-1} + A_0^{-1}V(I - W^T A_0^{-1} V)^{-1} W^T A_0^{-1}. \tag{3.7}$$

Since A_0 can be easily inverted, the main thing is to invert $(I - W^T A_0^{-1} V)$. The matrix $(I - W^T A_0^{-1} V)$ is a 4×4 matrix and is given by

$$(I - W^T A_0^{-1} V) = \begin{bmatrix} a_{11} & 0 & a_{13} & a_{14} \\ 0 & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & 0 \\ a_{41} & a_{42} & 0 & a_{44} \end{bmatrix}, \tag{3.8}$$

where

$$a_{11} = a_{22} = \sum_{j=1}^{n} \frac{2\omega_j(\mu_j + 2\frac{\mu_j^2}{\sigma_t h_i})}{\sigma_t h_i(1 + 2\frac{\mu_j}{\sigma_t h_i} + 2\frac{\mu_j^2}{\sigma_t^2 h_i^2})}, \tag{3.9a}$$

$$a_{33} = a_{44} = \sum_{j=1}^{n} \frac{2\omega_j(\mu_j + 2\frac{\mu_j^2}{\sigma_t h_{i+1}})}{\sigma_t h_{i+1}(1 + 2\frac{\mu_j}{\sigma_t h_{i+1}} + 2\frac{\mu_j^2}{\sigma_t^2 h_{i+1}^2})}, \tag{3.9b}$$

$$a_{13} = -\sum_{j=1}^{n} \frac{2\omega_j(\frac{\mu_i^2}{\sigma_t^2 h_i^2} + \frac{\mu_j^3}{\sigma_t^3 h_i^2 h_{i+1}})}{(1 + 2\frac{\mu_j}{\sigma_t h_i} + 2\frac{\mu_j^2}{\sigma_t^2 h_i^2})(1 + 2\frac{\mu_j}{\sigma_t h_{i+1}} + 2\frac{\mu_j^2}{\sigma_t^2 h_{i+1}^2})}, \tag{3.9c}$$

$$a_{14} = -\sum_{j=1}^{n} \frac{2\omega_j \mu_i^3}{\sigma_t^3 h_i^2 h_{i+1}(1 + 2\frac{\mu_j}{\sigma_t h_i} + 2\frac{\mu_j^2}{\sigma_t^2 h_i^2})(1 + 2\frac{\mu_j}{\sigma_t h_{i+1}} + 2\frac{\mu_j^2}{\sigma_t^2 h_{i+1}^2})}, \tag{3.9d}$$

$$a_{23} = -\sum_{j=1}^{n} \frac{2\omega_j(\frac{\mu_j}{\sigma_t h_i} + \frac{\mu_j^2}{\sigma_t^2 h_i^2} + \frac{\mu_j^2(1+\frac{\mu_j}{\sigma_t h_i})}{\sigma_t^2 h_i h_{i+1}})}{(1 + 2\frac{\mu_j}{\sigma_t h_i} + 2\frac{\mu_j^2}{\sigma_t^2 h_i^2})(1 + 2\frac{\mu_j}{\sigma_t h_{i+1}} + 2\frac{\mu_j^2}{\sigma_t^2 h_{i+1}^2})}, \tag{3.9e}$$

$$a_{24} = -\sum_{j=1}^{n} \frac{2\omega_j \frac{\mu_j^2}{\sigma_t^2 h_i h_{i+1}}(1 + \frac{\mu_j}{\sigma_t h_i})}{(1 + 2\frac{\mu_j}{\sigma_t h_i} + 2\frac{\mu_j^2}{\sigma_t^2 h_i^2})(1 + 2\frac{\mu_j}{\sigma_t h_{i+1}} + 2\frac{\mu_j^2}{\sigma_t^2 h_{i+1}^2})}, \tag{3.9f}$$

$$a_{31} = -\sum_{j=1}^{n} \frac{2\omega_j \frac{\mu_j^2}{\sigma_t^2 h_i h_{i+1}}(1 + \frac{\mu_j}{\sigma_t h_{i+1}})}{(1 + 2\frac{\mu_j}{\sigma_t h_i} + 2\frac{\mu_j^2}{\sigma_t^2 h_i^2})(1 + 2\frac{\mu_j}{\sigma_t h_{i+1}} + 2\frac{\mu_j^2}{\sigma_t^2 h_{i+1}^2})}, \tag{3.9g}$$

$$a_{32} = -\sum_{j=1}^{n} \frac{2\omega_j(\frac{\mu_j}{\sigma_t h_{i+1}} + \frac{\mu_j^2}{\sigma_t^2 h_{i+1}^2} + \frac{\mu_j^2(1+\frac{\mu_j}{\sigma_t h_{i+1}})}{\sigma_t^2 h_i h_{i+1}})}{(1 + 2\frac{\mu_j}{\sigma_t h_i} + 2\frac{\mu_j^2}{\sigma_t^2 h_i^2})(1 + 2\frac{\mu_j}{\sigma_t h_{i+1}} + 2\frac{\mu_j^2}{\sigma_t^2 h_{i+1}^2})}, \tag{3.9h}$$

$$a_{41} = -\sum_{j=1}^{n} \frac{2\omega_j \mu_i^3}{\sigma_t^3 h_i h_{i+1}^2(1 + 2\frac{\mu_j}{\sigma_t h_i} + 2\frac{\mu_j^2}{\sigma_t^2 h_i^2})(1 + 2\frac{\mu_j}{\sigma_t h_{i+1}} + 2\frac{\mu_j^2}{\sigma_t^2 h_{i+1}^2})}, \tag{3.9i}$$

$$a_{42} = -\sum_{j=1}^{n} \frac{2\omega_j(\frac{\mu_i^2}{\sigma_t^2 h_{i+1}^2} + \frac{\mu_j^3}{\sigma_t^3 h_i h_{i+1}^2})}{(1 + 2\frac{\mu_j}{\sigma_t h_i} + 2\frac{\mu_j^2}{\sigma_t^2 h_i^2})(1 + 2\frac{\mu_j}{\sigma_t h_{i+1}} + 2\frac{\mu_j^2}{\sigma_t^2 h_{i+1}^2})}, \tag{3.9j}$$

Since $(I - W^T A_0^{-1} V)$ is a 4 by 4 matrix, its inversion can be obtained inexpensively and explicitly by a direct method. It is obvious that no matter how many moments are used in (2.1), the μ-line relaxation can be carried out with the same insignificant expense.

The next three theorems examine the properties of two-cell red-black block μ-line relaxation. In the first theorem we consider the S_2 problem, that is, one positive and one negative μ. We assume the number of cells m is an even number.

Theorem 1 *For S_2, the errors after two-cell red-black μ-line relaxation will be piecewise linear across two cells.*

Proof: Since we use red-black relaxation, we can just examine each pair of two cells individually. When there is only one angle in the positive direction and one in the negative direction, $R = \frac{1}{2}, B_k = \frac{\mu_1}{\sigma_t h_k}, k = 1, ..., m$. At cells $2i-1$ and $2i$, after a relaxation, the error equation (3.4) becomes

$$(3.10) \qquad \begin{bmatrix} \underline{e}^-_{2i-\frac{3}{2}} \\ \underline{e}^+_{2i-1} \\ \underline{e}^-_{2i-1} \\ \underline{e}^+_{2i-\frac{1}{2}} \\ \underline{e}^-_{2i-\frac{1}{2}} \\ \underline{e}^+_{2i} \\ \underline{e}^-_{2i} \\ \underline{e}^+_{2i+\frac{1}{2}} \end{bmatrix} = \begin{bmatrix} A_{2i-1} & -C_{2i-1} \\ -D_{2i} & A_{2i} \end{bmatrix}^{-1} \begin{bmatrix} 0 \\ B_{2i-1}\underline{e}^+_{2i-\frac{3}{2}} \\ 0 \\ 0 \\ 0 \\ 0 \\ B_{2i}\underline{e}^-_{2i+\frac{1}{2}} \\ 0 \end{bmatrix}.$$

By setting $n = 1$, the explicit form of $\begin{bmatrix} A_{2i-1} & -C_{2i-1} \\ -D_{2i} & A_{2i} \end{bmatrix}^{-1}$ can be easily obtained by Sherman-Morrison formula (3.7). After multiplication in (3.10), we have

$$(3.11) \qquad \begin{bmatrix} \underline{e}^-_{2i-\frac{3}{2}} \\ \underline{e}^+_{2i-1} \\ \underline{e}^-_{2i-1} \\ \underline{e}^+_{2i-\frac{1}{2}} \\ \underline{e}^-_{2i-\frac{1}{2}} \\ \underline{e}^+_{2i} \\ \underline{e}^-_{2i} \\ \underline{e}^+_{2i+\frac{1}{2}} \end{bmatrix} = \frac{\sigma_t h_{2i}}{2\mu_1 + \sigma_t h_{2i-1} + \sigma_t h_{2i}} \begin{bmatrix} 1 + \frac{h_{2i-1}}{h_{2i}} \\ 1 + \frac{h_{2i-1}}{2h_{2i}} + \frac{2\mu_1}{\sigma_t h_{2i}} \\ 1 + \frac{h_{2i-1}}{2h_{2i}} \\ 1 + \frac{2\mu_1}{\sigma_t h_{2i}} \\ 1 \\ \frac{1}{2} + \frac{2\mu_1}{\sigma_t h_{2i}} \\ \frac{1}{2} \\ \frac{2\mu_1}{\sigma_t h_{2i}} \end{bmatrix} \underline{e}^+_{2i-\frac{3}{2}}$$

$$
+\frac{\sigma_t h_{2i-1}}{2\mu_1+\sigma_t h_{2i-1}+\sigma_t h_{2i}}\begin{bmatrix} \frac{2\mu_1}{\sigma_t h_{2i-1}} \\ \frac{1}{2} \\ \frac{1}{2}+\frac{2\mu_1}{\sigma_t h_{2i-1}} \\ 1 \\ 1+\frac{2\mu_1}{\sigma_t h_{2i-1}} \\ 1+\frac{h_{2i}}{2h_{2i-1}} \\ 1+\frac{h_{2i}}{2h_{2i-1}}+\frac{2\mu_1}{\sigma_t h_{2i-1}} \\ 1+\frac{h_{2i}}{h_{2i-1}} \end{bmatrix} \underline{e}^-_{2i+\frac{1}{2}}.
$$

The errors fit the following linear formulas:

(3.12a)

$$
\underline{e}^-_k = \frac{\sigma_t h_{2i}}{2\mu_1+\sigma_t h_{2i-1}+\sigma_t h_{2i}}(1+\frac{h_{2i-1}}{h_{2i}}-\frac{\xi_k}{h_{2i}})\underline{e}^+_{2i-\frac{3}{2}}+\frac{\sigma_t h_{2i-1}}{2\mu_1+\sigma_t h_{2i-1}+\sigma_t h_{2i}}(\frac{2\mu_1}{\sigma_t h_{2i-1}}+\frac{\xi_k}{h_{2i-1}})\underline{e}^-_{2i+\frac{1}{2}},
$$

where $k = 2i-\frac{3}{2}, 2i-1, 2i-\frac{1}{2}, 2i$, and $\xi_k = 0, \frac{h_{2i-1}}{2}, h_{2i-1}, h_{2i-1}+\frac{h_{2i}}{2}$ correspondingly and

(3.12b)

$$
\underline{e}^+_k = \frac{\sigma_t h_{2i}}{2\mu_1+\sigma_t h_{2i-1}+\sigma_t h_{2i}}(1+\frac{h_{2i-1}}{h_{2i}}+\frac{2\mu_1}{\sigma_t h_{2i}}-\frac{\xi_k}{h_{2i}})\underline{e}^+_{2i-\frac{3}{2}}+\frac{\sigma_t h_{2i-1}\xi_k}{h_{2i-1}(2\mu_1+\sigma_t h_{2i-1}+\sigma_t h_{2i})}\underline{e}^-_{2i+\frac{1}{2}},
$$

where $k = 2i-1, 2i-\frac{1}{2}, 2i, 2i+\frac{1}{2}$, and $\xi_k = \frac{h_{2i-1}}{2}, h_{2i-1}, h_{2i-1}+\frac{h_{2i}}{2}, h_{2i-1}+h_{2i}$ correspondingly. Since after relaxation the errors fit linear formulas at both the positive and negative angle direction, the errors are linearly distributed across two cells. $\square$

Let us define a new cell i on a coarser grid $\hat{h}$ whose length $\hat{h}_i$ is $\hat{h}_i = h_{2i-1}+h_{2i}$ and it covers cells $2i-1$ and cell $2i$. We can find values at the cell center and the edges of cell i of grid $\hat{h}$ by letting $\xi_k = 0, \xi_k = \frac{h_{2i-1}+h_{2i}}{2}$ and $\xi_k = h_{2i-1}+h_{2i}$ in (3.12). Thus, we have

(3.13a)

$$
\underline{e}^{-(\hat{h})}_{i-\frac{1}{2}} = \frac{\sigma_t h_{2i}}{2\mu_1+\sigma_t h_{2i-1}+\sigma_t h_{2i}}(1+\frac{h_{2i-1}}{h_{2i}})\underline{e}^+_{2i-\frac{3}{2}}+\frac{h_{2i-1}}{h_{2i-1}(2\mu_1+\sigma_t h_{2i-1}+\sigma_t h_{2i})}\underline{e}^-_{2i+\frac{1}{2}},
$$

(3.13b)

$$
\underline{e}^{+(\hat{h})}_{i} = \frac{\sigma_t h_{2i}}{2\mu_1+\sigma_t h_{2i-1}+\sigma_t h_{2i}}(\frac{1}{2}+\frac{h_{2i-1}}{2h_{2i}}+\frac{2\mu_1}{\sigma_t h_{2i}})\underline{e}^+_{2i-\frac{3}{2}}+\frac{\sigma_t h_{2i-1}}{2\mu_1+\sigma_t h_{2i-1}+\sigma_t h_{2i}}(\frac{1}{2}+\frac{h_{2i}}{2h_{2i-1}})\underline{e}^-_{2i+\frac{1}{2}},
$$

(3.13c)

$$
\underline{e}^{-(\hat{h})}_{i} = \frac{\sigma_t h_{2i}}{2\mu_1+\sigma_t h_{2i-1}+\sigma_t h_{2i}}(\frac{1}{2}+\frac{h_{2i-1}}{2h_{2i}})\underline{e}^+_{2i-\frac{3}{2}}+\frac{\sigma_t h_{2i-1}}{2\mu_1+\sigma_t h_{2i-1}+\sigma_t h_{2i}}(\frac{1}{2}+\frac{h_{2i}}{2h_{2i-1}}+\frac{2\mu_1}{\sigma_t h_{2i-1}})\underline{e}^-_{2i+\frac{1}{2}},
$$

(3.13d)

$$
\underline{e}^{+(\hat{h})}_{i+\frac{1}{2}} = \frac{\sigma_t h_{2i}}{2\mu_1+\sigma_t h_{2i-1}+\sigma_t h_{2i}}\frac{2\mu_1}{\sigma_t h_{2i}}\underline{e}^+_{2i-\frac{3}{2}}+\frac{\sigma_t h_{2i-1}}{2\mu_1+\sigma_t h_{2i-1}+\sigma_t h_{2i}}(1+\frac{h_{2i}}{h_{2i-1}})\underline{e}^-_{2i+\frac{1}{2}},
$$

We can then write (3.11) in terms of $\underline{e}_{i-\frac{1}{2}}^{-(\hat{h})}, \underline{e}_{i}^{+(\hat{h})}, \underline{e}_{i}^{-(\hat{h})}, \underline{e}_{i+\frac{1}{2}}^{+(\hat{h})}$ as

$$(3.14)\qquad \begin{bmatrix} \underline{e}_{2i-\frac{3}{2}}^{-} \\ \underline{e}_{2i-1}^{+} \\ \underline{e}_{2i-1}^{-} \\ \underline{e}_{2i-\frac{1}{2}}^{+} \\ \underline{e}_{2i-\frac{1}{2}}^{-} \\ \underline{e}_{2i}^{+} \\ \underline{e}_{2i}^{-} \\ \underline{e}_{2i+\frac{1}{2}}^{+} \end{bmatrix} = \begin{bmatrix} I & 0 & 0 & 0 \\ 0 & \frac{2h_{2i}+h_{2i-1}}{h_{2i}+h_{2i-1}}I & 0 & \frac{-h_{2i}}{h_{2i}+h_{2i-1}}I \\ \frac{h_{2i}}{h_{2i}+h_{2i-1}}I & 0 & \frac{h_{2i-1}}{h_{2i}+h_{2i-1}}I & 0 \\ 0 & \frac{2h_{2i}}{h_{2i}+h_{2i-1}}I & 0 & \frac{h_{2i-1}-h_{2i}}{h_{2i}+h_{2i-1}}I \\ \frac{h_{2i}-h_{2i-1}}{h_{2i}+h_{2i-1}}I & 0 & \frac{2h_{2i-1}}{h_{2i}+h_{2i-1}}I & 0 \\ 0 & \frac{h_{2i}}{h_{2i}+h_{2i-1}}I & 0 & \frac{h_{2i-1}}{h_{2i}+h_{2i-1}}I \\ \frac{-h_{2i-1}}{h_{2i}+h_{2i-1}}I & 0 & \frac{2h_{2i-1}+h_{2i}}{h_{2i}+h_{2i-1}}I & 0 \\ 0 & 0 & 0 & I \end{bmatrix} \begin{bmatrix} \underline{e}_{i-\frac{1}{2}}^{-(\hat{h})} \\ \underline{e}_{i}^{+(\hat{h})} \\ \underline{e}_{i}^{-(\hat{h})} \\ \underline{e}_{i+\frac{1}{2}}^{+(\hat{h})} \end{bmatrix}.$$

Let

$$(3.15a)\qquad T_{1,i} = \begin{bmatrix} I & 0 & 0 & 0 \\ 0 & \frac{2h_{i+1}+h_i}{h_{i+1}+h_i}I & 0 & \frac{-h_{i+1}}{h_{i+1}+h_i}I \\ \frac{h_{i+1}}{h_{i+1}+h_i}I & 0 & 0\frac{h_i}{h_{i+1}+h_i}I & 0 \\ 0 & \frac{2h_{i+1}}{h_{i+1}+h_i}I & 0 & \frac{h_i-h_{i+1}}{h_{i+1}+h_i}I \end{bmatrix},$$

$$(3.15b)\qquad T_{2,i} = \begin{bmatrix} \frac{h_{i+1}-h_i}{h_{i+1}+h_i}I & 0 & \frac{2h_i}{h_{i+1}+h_i}I & 0 \\ 0 & \frac{h_{i+1}}{h_{i+1}+h_i}I & 0 & \frac{h_i}{h_{i+1}+h_i}I \\ \frac{-h_i}{h_{i+1}+h_i}I & 0 & \frac{2h_i+h_{i+1}}{h_{i+1}+h_i}I & 0 \\ 0 & 0 & 0 & I \end{bmatrix}.$$

Then after the relaxation, the errors can be written as

$$(3.16)\qquad \begin{bmatrix} \underline{e}_{\frac{1}{2}}^{-} \\ \underline{e}_{1}^{+} \\ \underline{e}_{1}^{-} \\ \underline{e}_{\frac{3}{2}}^{+} \\ \vdots \\ \underline{e}_{m-\frac{1}{2}}^{-} \\ \underline{e}_{m}^{+} \\ \underline{e}_{m}^{-} \\ \underline{e}_{m+\frac{1}{2}}^{+} \end{bmatrix} = \begin{bmatrix} T_{1,1} & & & & \\ T_{2,1} & & & & \\ & T_{1,3} & & & \\ & T_{2,3} & & & \\ & & \ddots & & \\ & & \ddots & & \\ & & & & T_{1,m-1} \\ & & & & T_{2,m-1} \end{bmatrix} \begin{bmatrix} \underline{e}_{\frac{1}{2}}^{-(\hat{h})} \\ \underline{e}_{1}^{+(\hat{h})} \\ \underline{e}_{1}^{-(\hat{h})} \\ \underline{e}_{\frac{3}{2}}^{+(\hat{h})} \\ \vdots \\ \underline{e}_{\frac{m}{2}-\frac{1}{2}}^{-(\hat{h})} \\ \underline{e}_{\frac{m}{2}}^{+(\hat{h})} \\ \underline{e}_{\frac{m}{2}}^{-(\hat{h})} \\ \underline{e}_{\frac{m}{2}+\frac{1}{2}}^{+(\hat{h})} \end{bmatrix}.$$

We write (3.16) in multigrid notation as

$$(3.17)\qquad \underline{e}^{h} = I_{2h}^{h}\underline{e}^{2h},$$

where

$$
I_{2h}^{h} = \begin{bmatrix} T_{1,1} & & & & \\ T_{2,1} & & & & \\ & T_{1,3} & & & \\ & T_{2,3} & & & \\ & & \ddots & & \\ & & \ddots & & \\ & & & & T_{1,m-1} \\ & & & & T_{2,m-1} \end{bmatrix}. \tag{3.18}
$$

We use the notation $2h$ to indicate a coarse grid although our grid is not uniform. In the nonuniform case, the mesh size $\hat{h}_i$ at cell i on the coarse grid is $h_{2i-1} + h_{2i}$ of the fine grid instead of being twice as much as that in fine grid. In the next section, we will see that (3.16) means the errors after relaxation are in the range of interpolation.

In the multi-angle case, the errors will no longer be exactly linearly distributed across two cells after two-cell red-black block μ-line relaxation. But, we can examine the property of the relaxation in two extreme cases. These are when $\max\limits_{k=(1,\ldots,m)} (\sigma_t h_k) \ll 1$ and $\min\limits_{k=(1,\ldots,m)} (\sigma_t h_k) \gg 1$. The next two theorems will give Taylor expansion forms of the errors after the relaxation.

Theorem 2 *When* $\max(\sigma_t h_{2i-1}, \sigma_t h_{2i}) \ll 1$, *the errors after two-cell red-black block* μ*-line relaxation will be piecewise linear across two cells up to the accuracy of* $O(\max(\sigma_t^2 h_{2i-1}^2, \sigma_t^2 h_{2i}^2))$.

Proof: We can write $\begin{bmatrix} A_{2i-1} & -C_{2i-1} \\ -D_{2i} & A_{2i} \end{bmatrix}$ in (3.10) as

$$
\begin{bmatrix} A_{2i-1} & -C_{2i-1} \\ -D_{2i} & A_{2i} \end{bmatrix} = H_0 + H_1 \tag{3.19}
$$

with

(3.20a)

$$
H_0 = \begin{bmatrix}
\frac{2}{\sigma_t h_{2i-1}} M & 0 & -\frac{2}{\sigma_t h_{2i-1}} M & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & \frac{1}{\sigma_t h_{2i-1}} M & 0 & 0 & 0 & 0 \\
\frac{1}{\sigma_t h_{2i-1}} & 0 & 0 & 0 & -\frac{1}{\sigma_t h_{2i-1}} M & 0 & 0 & 0 \\
0 & -\frac{2}{\sigma_t h_{2i-1}} M & 0 & \frac{2}{\sigma_t h_{2i-1}} M & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & \frac{2}{\sigma_t h_{2i}} M & 0 & -\frac{2}{\sigma_t h_{2i}} M & 0 \\
0 & 0 & 0 & -\frac{1}{\sigma_t h_{2i}} M & 0 & 0 & 0 & \frac{1}{\sigma_t h_{2i}} M \\
0 & 0 & 0 & 0 & \frac{1}{\sigma_t h_{2i}} M & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & -\frac{2}{\sigma_t h_{2i}} M & 0 & \frac{2}{\sigma_t h_{2i}} M
\end{bmatrix},
$$

$$
H_1 = \begin{bmatrix} I-R & -2R & 0 & R & 0 & 0 & 0 & 0 \\ 0 & I-R & -R & 0 & 0 & 0 & 0 & 0 \\ 0 & -R & I-R & 0 & 0 & 0 & 0 & 0 \\ R & 0 & -2R & I-R & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & I-R & -2R & 0 & R \\ 0 & 0 & 0 & 0 & 0 & I-R & -R & 0 \\ 0 & 0 & 0 & 0 & 0 & -R & I-R & 0 \\ 0 & 0 & 0 & 0 & R & 0 & -2R & I-R \end{bmatrix}, \tag{3.20b}
$$

and

$$
M = \sigma_t h_i B_i = \begin{bmatrix} \mu_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \mu_n \end{bmatrix}. \tag{3.20c}
$$

Then

$$
\begin{bmatrix} A_{2i-1} & -C_{2i-1} \\ -D_{2i} & A_{2i} \end{bmatrix}^{-1} = (I + H_0^{-1} H_1)^{-1} H_0^{-1}. \tag{3.21}
$$

It is trivial to obtain

(3.22)

$$
H_0^{-1} = \begin{bmatrix} 0 & 0 & \sigma_t h_{2i-1} M^{-1} & 0 & 0 & 0 & \sigma_t h_{2i} M^{-1} & 0 \\ 0 & \sigma_t h_{2i-1} M^{-1} & 0 & -\frac{\sigma_t h_{2i-1}}{2} M^{-1} & 0 & 0 & 0 & 0 \\ -\frac{\sigma_t h_{2i-1}}{2} M^{-1} & 0 & \sigma_t h_{2i-1} M^{-1} & 0 & 0 & 0 & \sigma_t h_{2i} M^{-1} & 0 \\ 0 & \sigma_t h_{2i-1} M^{-1} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \sigma_t h_{2i} M^{-1} & 0 \\ 0 & \sigma_t h_{2i-1} M^{-1} & 0 & 0 & 0 & \sigma_t h_{2i} M^{-1} & 0 & -\frac{\sigma_t h_{2i}}{2} M^{-1} \\ 0 & 0 & 0 & 0 & -\frac{\sigma_t h_{2i}}{2} M^{-1} & 0 & \sigma_t h_{2i} M^{-1} & 0 \\ 0 & \sigma_t h_{2i-1} M^{-1} & 0 & 0 & 0 & \sigma_t h_{2i} M^{-1} & 0 & 0 \end{bmatrix}.
$$

When $\sigma_t h_{2i-1}, \sigma_t h_{2i} \ll 1$, $\|H_0^{-1}\| \ll 1$. Since H_1 is a constant matrix, $\|H_0^{-1} H_1\| \ll 1$. So $I + H_0^{-1} H_1$ is invertable and

$$
(I + H_0^{-1} H_1)^{-1} = I - H_0^{-1} H_1 + O(max(\sigma_t^2 h_{2i-1}^2, \sigma_t^2 h_{2i}^2)). \tag{3.23}
$$

Substituting (3.23) into (3.21), we have

$$
\begin{bmatrix} A_{2i-1} & -C_{2i-1} \\ -D_{2i} & A_{2i} \end{bmatrix}^{-1} = H_0^{-1} - H_0^{-1} H_1 H_0^{-1} + O(max(\sigma_t^3 h_{2i-1}^3, \sigma_t^3 h_{2i}^3)). \tag{3.24}
$$

When we carry out the arithmetic in (3.24) by using $(3.20b)$ and (3.22) and substitute the resulting formula into (3.10), we have the errors after relaxation as

$$(3.25)\qquad \begin{bmatrix} \underline{e}^-_{2i-\frac{3}{2}} \\ \underline{e}^+_{2i-1} \\ \underline{e}^-_{2i-1} \\ \underline{e}^+_{2i-\frac{1}{2}} \\ \underline{e}^-_{2i-\frac{1}{2}} \\ \underline{e}^+_{2i} \\ \underline{e}^-_{2i} \\ \underline{e}^+_{2i+\frac{1}{2}} \end{bmatrix} = \begin{bmatrix} I \\ 0 \\ I \\ 0 \\ I \\ 0 \\ I \\ 0 \end{bmatrix} \underline{e}^+_{2i-\frac{3}{2}} + \sigma_t h_{2i-1} \begin{bmatrix} (1+\frac{h_{2i}}{h_{2i-1}})M^{-1}(I-R) \\ -\frac{1}{2}M^{-1}R \\ (\frac{1}{2}+\frac{h_{2i}}{h_{2i-1}})M^{-1}(I-R) \\ -M^{-1}R \\ \frac{h_{2i}}{h_{2i-1}}M^{-1}(I-R) \\ -(1+\frac{h_{2i}}{2h_{2i-1}})M^{-1}R \\ \frac{h_{2i}}{2h_{2i-1}}M^{-1}(I-R) \\ -(1+\frac{h_{2i}}{h_{2i-1}})M^{-1}R \end{bmatrix} \underline{e}^+_{2i-\frac{3}{2}}.$$

$$+ \begin{bmatrix} 0 \\ I \\ 0 \\ I \\ 0 \\ I \\ 0 \\ I \end{bmatrix} \underline{e}^-_{2i+\frac{1}{2}} + \sigma_t h_{2i} \begin{bmatrix} -(1+\frac{h_{2i-1}}{h_{2i}})M^{-1}R \\ \frac{h_{2i-1}}{2h_{2i}}M^{-1}(I-R) \\ -(1+\frac{h_{2i-1}}{2h_{2i}})M^{-1}R \\ \frac{h_{2i-1}}{h_{2i}}M^{-1}(I-R) \\ -M^{-1}R \\ (\frac{1}{2}+\frac{h_{2i-1}}{h_{2i}})M^{-1}(I-R) \\ -\frac{1}{2}M^{-1}R \\ (1+\frac{h_{2i-1}}{h_{2i}})M^{-1}(I-R) \end{bmatrix} \underline{e}^-_{2i+\frac{1}{2}} + O(max(\sigma_t^2 h_{2i-1}^2, \sigma_t^2 h_{2i}^2))$$

When we define the same cell i on grid $\hat{h}$ as in Theorem 1, we have

$$(3.26a)\quad \underline{e}^{-(\hat{h})}_{i-\frac{1}{2}} = (1-\sigma_t h_{2i-1}(1+\frac{h_{2i}}{h_{2i-1}}))M^{-1}(I-R)\underline{e}^+_{2i-\frac{3}{2}} + \sigma_t h_{2i}(1+\frac{h_{2i-1}}{h_{2i}})M^{-1}R\underline{e}^-_{2i+\frac{1}{2}},$$

(3.26b)

$$\underline{e}^{+(\hat{h})}_i = \sigma_t h_{2i-1}(\frac{1}{2}+\frac{h_{2i}}{2h_{2i-1}})M^{-1}R\underline{e}^+_{2i-\frac{3}{2}} + (1-\sigma_t h_{2i}(\frac{1}{2}+\frac{h_{2i-1}}{2h_{2i}}))M^{-1}(I-R)\underline{e}^-_{2i+\frac{1}{2}},$$

(3.26c)

$$\underline{e}^{-(\hat{h})}_i = (1-\sigma_t h_{2i-1}(\frac{1}{2}+\frac{h_{2i}}{2h_{2i-1}}))M^{-1}(I-R)\underline{e}^+_{2i-\frac{3}{2}} + \sigma_t h_{2i}(\frac{1}{2}+\frac{2h_{2i-1}}{h_{2i}})M^{-1}R\underline{e}^-_{2i+\frac{1}{2}},$$

$$(3.26d)\quad \underline{e}^{+(\hat{h})}_{i+\frac{1}{2}} = \sigma_t h_{2i-1}(1+\frac{h_{2i}}{h_{2i-1}})M^{-1}R\underline{e}^+_{2i-\frac{3}{2}} + (1-\sigma_t h_{2i}(1+\frac{h_{2i-1}}{2h_{2i}}))M^{-1}(I-R)\underline{e}^-_{2i+\frac{1}{2}}.$$

We can write (3.25) in terms of $\underline{e}_{i-\frac{1}{2}}^{-(\hat{h})}, \underline{e}_i^{+(\hat{h})}, \underline{e}_i^{-(\hat{h})}, \underline{e}_{i+\frac{1}{2}}^{+(\hat{h})}$ as

$$
(3.27) \qquad \begin{bmatrix} \underline{e}_{2i-\frac{3}{2}}^{-} \\ \underline{e}_{2i-1}^{+} \\ \underline{e}_{2i-1}^{-} \\ \underline{e}_{2i-\frac{1}{2}}^{+} \\ \underline{e}_{2i-\frac{1}{2}}^{-} \\ \underline{e}_{2i}^{+} \\ \underline{e}_{2i}^{-} \\ \underline{e}_{2i+\frac{1}{2}}^{+} \end{bmatrix} = \begin{bmatrix} T_{1,2i-1} \\ T_{2,2i-1} \end{bmatrix} \begin{bmatrix} \underline{e}_{i-\frac{1}{2}}^{-(\hat{h})} \\ \underline{e}_i^{+(\hat{h})} \\ \underline{e}_i^{-(\hat{h})} \\ \underline{e}_{i+\frac{1}{2}}^{+(\hat{h})} \end{bmatrix} + O(max(\sigma_t^2 h_{2i-1}^2, \sigma_t^2 h_{2i}^2)).
$$

Thus, the errors are piecewise linear across the two cells $2i-1$ and $2i$ up to the accuracy of $O(max(\sigma_t^2 h_{2i-1}^2, \sigma_t^2 h_{2i}^2))$. $\square$

We now focus our attention on the optically dense or thick limit.

Lemma 1 . *Suppose* $(\sigma_t h_i) \gg 1$*; then, the asymptotic inversion of* A_i *defined by* $(2.18a)$ *can be expanded as*

$$
(3.28) \qquad A_i^{-1} = \frac{\sigma_t h_i}{2c_0} \begin{bmatrix} R & 2R & 0 & -R \\ 0 & R & R & 0 \\ 0 & R & R & 0 \\ -R & 0 & 2R & R \end{bmatrix} + I
$$

$$
+\frac{1}{2c_0} \begin{bmatrix} -2(MR+RM) & -2(MR+RM) & 2(MR+RM) & 2MR \\ (MR-RM) & 0 & -2MR & -(MR+RM) \\ -(MR+RM) & -2MR & 0 & (MR-RM) \\ 2MR & 2(MR+RM) & -2(MR+RM) & -2(MR+RM) \end{bmatrix} + O(\frac{1}{\sigma_t h_i}).
$$

Proof: Write

$$
(3.29) \qquad A_i = \begin{bmatrix} I+2B_i-R & -2R & -2B_i & R \\ 0 & I-R & -R & B_i \\ B_i & -R & I-R & 0 \\ R & -2B_i & -2R & I+2B_i-R \end{bmatrix}
$$

$$
= \begin{bmatrix} I+2B_i & 0 & -2B_i & 0 \\ 0 & I & 0 & B_i \\ B_i & 0 & I & 0 \\ 0 & -2B_i & 0 & I+2B_i \end{bmatrix} - \begin{bmatrix} 2\underline{1} & 0 \\ \underline{1} & \underline{1} \\ \underline{1} & \underline{1} \\ 0 & 2\underline{1} \end{bmatrix} \begin{bmatrix} \frac{1}{2}\underline{w}^T & \underline{w}^T & 0 & -\frac{1}{2}\underline{w}^T \\ -\frac{1}{2}\underline{w}^T & 0 & \underline{w}^T & \frac{1}{2}\underline{w}^T \end{bmatrix},
$$

where $\underline{1} = (1,1,...,1)^T$ and $\underline{w}^T = (\omega_1, \omega_2, ..., \omega_n)$. The first term is easily inverted while the second term is a rank two matrix. If we denote this as

$$
(3.30) \qquad A_i = A_0 - VW^T,
$$

with A_0 denoting the first term and VW^T the second. Then the Sherman-Morrison formulas yield

$$A_i^{-1} = A_0^{-1} + A_0^{-1} V (I - W^T A_0^{-1} V)^{-1} W^T A_0^{-1}. \tag{3.31}$$

A_0^{-1} is readily given by

$$A_0 = (I + 2B_i + 2B_i^2)^{-1} \begin{bmatrix} I & 0 & 2B_i & 0 \\ 0 & I+2B_i & 0 & -B_i \\ -B_i & 0 & I+2B_i & 0 \\ 0 & 2B_i & 0 & I \end{bmatrix}. \tag{3.32}$$

The matrix $(I - W^T A_0^{-1} V)$ is 2×2 and its inverse can be obtained as

$$(I - W^T A_0^{-1} V)^{-1} = \xi_i \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \tag{3.33}$$

where

$$\xi_i = \frac{1}{1 - \zeta_i}, \quad \zeta_i = \sum_{j=1}^{N} \Big(\frac{2\omega_j (1 + \frac{\mu_j}{\sigma_t h_i})}{1 + 2\frac{\mu_j}{\sigma_t h_i} + 2(\frac{\mu_j}{\sigma_t h_i})^2} \Big). \tag{3.34}$$

In the optically dense media case, we expand ξ_i as

$$\xi_i = \frac{\sigma_t h_i}{2c_0} + O(\frac{1}{\sigma_t h_i}), \tag{3.35}$$

and

$$(I + 2B_i + 2B_i^2)^{-1} = I - \frac{2}{\sigma_t h_i} M + \frac{2}{\sigma_t^2 h_i^2} M^2 + O(\frac{1}{\sigma_t^3 h_i^3}). \tag{3.36}$$

Substituting (3.32),(3.33), (3.35) and (3.36) into (3.31), we have (3.28). $\square$

Theorem 3 . *Suppose* $\min_{k=(2i-1,2i)} (\sigma_t h_k) \gg 1$*; then, the errors after the two cell red-black* μ*-line relaxation will be piecewise linear up to the accuracy of* $O(\max(\frac{1}{\sigma_t h_{2i-1}}, \frac{1}{\sigma_t h_{2i}}))$.

Proof: When $\min_{k=(2i-1,2i)} (\sigma_t h_k) \gg 1$, we need to expand $\begin{bmatrix} A_{2i-1} & -C_{2i-1} \\ -D_{2i} & A_{2i} \end{bmatrix}^{-1}$. Write

$$\begin{bmatrix} A_{2i-1} & -C_{2i-1} \\ -D_{2i} & A_{2i} \end{bmatrix} = \begin{bmatrix} A_{2i-1} & 0 \\ -D_{2i} & A_{2i} \end{bmatrix} - \begin{bmatrix} 0 & C_{2i-1} \\ 0 & 0 \end{bmatrix} = A_0 - VW^T, \tag{3.37}$$

where

$$VW^T = \begin{bmatrix} 0 & C_{2i-1} \\ 0 & 0 \end{bmatrix}, \tag{3.38}$$

with $V = (0,0,I,0,0,0,0,0)^T$ and $W^T = (0,0,0,0,B_{2i-1},0,0,0)$. Its inversion can be obtained by Sherman-Morrison formula (3.31). We first note that

$$A_0^{-1} = \begin{bmatrix} A_{2i-1}^{-1} & 0 \\ A_{2i-1}^{-1} D_{2i} A_{2i}^{-1} & A_{2i}^{-1} \end{bmatrix}. \tag{3.39}$$

We can expand A_0^{-1} by using Lemma 1. By substituting the expansion of A_0^{-1} into $(I - W^T A_0^{-1} V)$, we can then expand $(I - W^T A_0^{-1} V)^{-1}$. The explicit expansion of $\begin{bmatrix} A_{2i-1} & -C_{2i-1} \\ -D_{2i} & A_{2i} \end{bmatrix}^{-1}$ is omitted here because of its complexity. After the expansion, the error equation (3.10) at cells $2i-1$ and $2i$ after two-cell μ-line relaxation can be expanded as

(3.40)

$$\begin{bmatrix} \underline{e}_{2i-\frac{3}{2}}^{-} \\ \underline{e}_{2i-1}^{+} \\ \underline{e}_{2i-1}^{-} \\ \underline{e}_{2i-\frac{1}{2}}^{+} \\ \underline{e}_{2i-\frac{1}{2}}^{-} \\ \underline{e}_{2i}^{+} \\ \underline{e}_{2i}^{-} \\ \underline{e}_{2i+\frac{1}{2}}^{+} \end{bmatrix} = \begin{bmatrix} R \\ (\frac{1}{2} + \frac{h_{2i-1}}{2(h_{2i-1}+h_{2i})})R \\ (\frac{1}{2} + \frac{h_{2i-1}}{2(h_{2i-1}+h_{2i})})R \\ \frac{h_{2i-1}}{(h_{2i-1}+h_{2i})}R \\ \frac{h_{2i-1}}{(h_{2i-1}+h_{2i})}R \\ \frac{h_{2i-1}}{2(h_{2i-1}+h_{2i})}R \\ \frac{h_{2i-1}}{2(h_{2i-1}+h_{2i})}R \\ 0 \end{bmatrix} \underline{e}_{2i-\frac{3}{2}}^{+} + \begin{bmatrix} 0 \\ \frac{h_{2i-1}}{2(h_{2i-1}+h_{2i})}R \\ \frac{h_{2i-1}}{2(h_{2i-1}+h_{2i})}R \\ \frac{h_{2i-1}}{(h_{2i-1}+h_{2i})}R \\ \frac{h_{2i-1}}{(h_{2i-1}+h_{2i})}R \\ (\frac{1}{2} + \frac{h_{2i-1}}{2(h_{2i-1}+h_{2i})})R \\ (\frac{1}{2} + \frac{h_{2i-1}}{2(h_{2i-1}+h_{2i})})R \\ R \end{bmatrix} \underline{e}_{2i+\frac{1}{2}}^{-} + O(max(\frac{1}{\sigma_t h_{2i-1}}, \frac{1}{\sigma_t h_{2i}})).$$

As in Theorem 1 and 2 we define the same cell i on grid $\hat{h}$ and obtain

$$\underline{e}_{i-\frac{1}{2}}^{-(\hat{h})} = R\underline{e}_{2i-\frac{3}{2}}^{+}, \tag{3.41a}$$

$$\underline{e}_{i}^{+(\hat{h})} = \frac{1}{2}R\underline{e}_{2i-\frac{3}{2}}^{+} + \frac{1}{2}R\underline{e}_{2i+\frac{1}{2}}^{-}, \tag{3.41b}$$

$$\underline{e}_{i}^{-(\hat{h})} = \frac{1}{2}R\underline{e}_{2i-\frac{3}{2}}^{+} + \frac{1}{2}R\underline{e}_{2i+\frac{1}{2}}^{-}, \tag{3.41c}$$

$$\underline{e}_{i+\frac{1}{2}}^{+(\hat{h})} = R\underline{e}_{2i+\frac{1}{2}}^{-}, \tag{3.41d}$$

Then (3.40) can be written as

$$\begin{bmatrix} \underline{e}_{2i-\frac{3}{2}}^{-} \\ \underline{e}_{2i-1}^{+} \\ \underline{e}_{2i-1}^{-} \\ \underline{e}_{2i-\frac{1}{2}}^{+} \\ \underline{e}_{2i-\frac{1}{2}}^{-} \\ \underline{e}_{2i}^{+} \\ \underline{e}_{2i}^{-} \\ \underline{e}_{2i+\frac{1}{2}}^{+} \end{bmatrix} = \begin{bmatrix} T_{1,2i-1} \\ T_{2,2i-1} \end{bmatrix} \begin{bmatrix} \underline{e}_{i-\frac{1}{2}}^{-(\hat{h})} \\ \underline{e}_{i}^{+(\hat{h})} \\ \underline{e}_{i}^{-(\hat{h})} \\ \underline{e}_{i+\frac{1}{2}}^{+(\hat{h})} \end{bmatrix} + O(max(\frac{1}{\sigma_t h_{2i-1}}, \frac{1}{\sigma_t h_{2i}})). \tag{3.42}$$

Therefore the leading terms in the errors after relaxation are piecewise linear across the two cells $2i-1$ and $2i$ up to the accuracy of $O(\max(\frac{1}{\sigma_t h_{2i-1}}, \frac{1}{\sigma_t h_{2i}}))$. $\square$

The property of the leading error being piecewise linear across two cells will be very important in our multigrid algorithm. In the next section, we will introduce the notations and strategy of the multigrid solution technique.

§4 Multigrid Solution Technique

Define A^h be the fine grid operator, A^{2h} the coarse grid operator and I_{2h}^h to be interpolation operator which is give by (3.18). I_h^{2h} is defined as a restriction operator which will transfer residuals from the fine grid to coarse grid. We will give the exact form of I_h^{2h} later. Then, a three-level CS(Correction Scheme)(Brandt[4])$V(\nu_1,\nu_2)$ multigrid method proceeds as follows:

i) relax ν_1 times on $A^h u^h = f^h$
ii) calculate residual $r^h = f^h - A^h u^h$ and transfer residual to grid $2h$ $\quad f^{2h} = I_h^{2h} r^h$
iii) relax ν_1 times on $A^{2h}u^{2h} = f^{2h}$
iv) calculate residual $r^{2h} = f^{2h} - A^{2h}u^{2h}$ and transfer residual to grid $4h$ $f^{4h} = I_{2h}^{4h} r^{2h}$
v) relax $\nu_1+\nu_2$ times on $A^{4h}u^{4h} = f^{4h}$ or solve it exactly
vi) replace $u^{2h} \leftarrow u^{2h} + I_{4h}^{2h} u^{4h}$ and relax ν_2 times on $A^{2h}u^{2h} = f^{2h}$
vii) replace $u^h \leftarrow u^h + I_{2h}^h u^{2h}$ and relax ν_2 times on $A^h u^h = f^h$

We need to define interpolation and restriction operators for the modified linear discontinuous scheme. By the same finite element principle as in derivation of MLD scheme, we can derive the restriction operator I_h^{2h} and interpolation operator I_{2h}^h.

From (2.17), the fine grid operator A^h is

$$A^h = \begin{bmatrix} A_1^h & -C_1^h & & & & \\ -D_2^h & A_2^h & -C_2^h & & & \\ & \ddots & \ddots & \ddots & & \\ & & -D_i^h & A_i^h & -C_i^h & \\ & & & \ddots & \ddots & \ddots \\ & & & & -D_m^h & A_m^h \end{bmatrix}, \tag{4.1}$$

where

$$A_i^h = \begin{bmatrix} I+2B_i^h - R & -2R & -2B_i^h & R \\ 0 & I-R & -R & B_i^h \\ B_i^h & -R & I-R & 0 \\ R & -2B_i^h & -2R & I+2B_i^h - R \end{bmatrix}, \tag{4.2a}$$

$$C_i^h = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ B_i^h & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, D_i^h = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & B_i^h \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, B_i^h = \frac{1}{\sigma_t h_i} \begin{bmatrix} \mu_1 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \mu_n \end{bmatrix}. \tag{4.2b}$$

The coarse grid operator A^{2h} is defined as

$$A^{2h} = I_h^{2h} A^h I_{2h}^h, \tag{4.3}$$

with restriction operator I_h^{2h} defined by

$$I_h^{2h} = \begin{bmatrix} S_{1,1} & S_{1,2} & & & & & \\ & & S_{3,1} & S_{3,2} & & & \\ & & & & \ddots & \ddots & \\ & & & & & S_{m-1,1} & S_{m-1,2} \end{bmatrix}, \tag{4.4}$$

$$S_{i,1} = \begin{bmatrix} I & 0 & 0 & 0 \\ 0 & \frac{h_i}{h_{i+1}+h_i} I & 0 & 0 \\ 0 & 0 & \frac{h_i}{h_{i+1}+h_i} I & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, S_{i,2} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & \frac{h_{i+1}}{h_{i+1}+h_i} I & 0 & 0 \\ 0 & 0 & \frac{h_{i+1}}{h_{i+1}+h_i} I & 0 \\ 0 & 0 & 0 & I \end{bmatrix}, \tag{4.5}$$

and interpolation operator I_{2h}^h defined by (3.18).

It is easy to show that A^{2h} correspond to the MLD scheme on the coarse grid, that is,

$$A^{2h} = \begin{bmatrix} A_1^{2h} & -C_1^{2h} & & & & & \\ -D_2^{2h} & A_2^{2h} & -C_2^{2h} & & & & \\ & \ddots & \ddots & \ddots & & & \\ & & -D_i^{2h} & A_i^{2h} & -C_i^{2h} & & \\ & & & \ddots & \ddots & \ddots & \\ & & & & & -D_{m/2}^{2h} & A_{m/2}^{2h} \end{bmatrix}, \tag{4.6}$$

where

$$A_i^{2h} = \begin{bmatrix} I + 2B_i^{2h} - R & -2R & -2B_i^{2h} & R \\ 0 & I - R & -R & B_i^{2h} \\ B_i^{2h} & -R & I - R & 0 \\ R & -2B_i^{2h} & -2R & I + 2B_i^{2h} - R \end{bmatrix}, \tag{4.7a}$$

(4.7b)

$$C_i^{2h} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ B_i^{2h} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \quad D_i^{2h} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & B_i^{2h} \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, B_i^{2h} = \frac{1}{\sigma_t \hat{h}_i} \begin{bmatrix} \mu_1 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \mu_n \end{bmatrix},$$

with $\hat{h}_i = h_{2i-1} + h_{2i}$ and $i = 1, \ldots, \frac{m}{2}$.

In the next subsection, we analyze the convergence factors of the multigrid algorithm defined above.

§4.1 Analysis of Convergence Factors of the Multigrid Algorithm

We start our analysis with the S_2 problem.

Theorem 4 *For the S_2 angular discretization, that is, when there is only one angle in the positive direction and one angle in the negative direction, the multigrid algorithm with two-cell red-black block μ-line relaxation will be an exact solver provided the coarsest grid is solved exactly.*

Proof: Define G^h be the two-cell red-black block μ-line relaxation matrix. The errors after the relaxation will be

$$\underline{e}_1^h = G^h \underline{e}_0^h, \tag{4.8}$$

where $\underline{e}_0^h$ and $\underline{e}_1^h$ are errors before and after relaxation respectively. From (3.16) in Theorem 1, the errors after the relaxation can be expressed as

$$\underline{e}_1^h = G^h \underline{e}_0^h = I_{2h}^h \underline{e}^{2h}, \tag{4.9}$$

where $\underline{e}^{2h}$ is defined on a coarse grid $\hat{h}$ with mesh size of cell i as $\hat{h}_i = h_{2i-1} + h_{2i}$ and its cell edges at the same spatial position of left edge of cell $2i-1$ and right edge of cell $2i$ of fine grid h. The error, $\underline{e}_2^h$, after one step of a two-grid multigrid $V(1,1)$ cycle can be written in matrix form as

$$\begin{aligned} \underline{e}_2^h &= G^h(G^h\underline{e}_0^h - I_{2h}^h (A^{2h})^{-1} I_h^{2h} A^h G^h \underline{e}_0^h) \\ &= G^h(\underline{e}_1^h - I_{2h}^h (A^{2h})^{-1} I_h^{2h} A^h \underline{e}_1^h). \end{aligned} \tag{4.10}$$

Substituting (4.9) into (4.10), we have

$$\underline{e}_2^h = G^h(I_{2h}^h \underline{e}^{2h} - I_{2h}^h (A^{2h})^{-1} I_h^{2h} A^h I_{2h}^h \underline{e}^{2h}). \tag{4.11}$$

Note that the coarse grid operator A^{2h} is defined by (4.3). So (4.11) becomes

$$\underline{e}_2^h = G^h(I_{2h}^h \underline{e}^{2h} - I_{2h}^h (A^{2h})^{-1} A^{2h} \underline{e}^{2h}) = 0. \tag{4.12}$$

This means that the errors will be completely eliminated by a multigrid V-cycle provided the coarse grid $2h$ is solved exactly. By induction, the coarse grid can be solved exactly if the coarsest grid solution is exact. The coarsest grid in our algorithm contains only two cells and can be exactly solved by two-cell block μ-line relaxation. □

For S_N with $N > 2$, the multigrid algorithm will no longer be an exact solver. However, we can establish upper bounds on the convergence factors of the multigrid algorithm when $\max\limits_{k=(1,\ldots,m)} (\sigma_t h_k) \ll 1$ and $\min\limits_{k=(1,\ldots,m)} (\sigma_t h_k) \gg 1$.

In the next two lemmas, we present upper bounds on the norm of the inversion of the operator A^{2h}. We use the notation $||.||_{1,\infty}$ to indicate that the bounds hold for both the L_1 norm and the L_∞ norm.

Lemma 2 *When* $\max(\sigma_t \hat{h}_i, \sigma_t \hat{h}_{i+1}) \ll 1$,

$$(4.13) \qquad || \begin{bmatrix} A_i^{2h} & -C_i^{2h} \\ -D_{i+1}^{2h} & A_{i+1}^{2h} \end{bmatrix}^{-1} ||_{1,\infty} \leq c_1 \max(\sigma_t \hat{h}_i, \sigma_t \hat{h}_{i+1}),$$

where $\hat{h}_i$ and $\hat{h}_{i+1}$ are the coarse grid cell sizes and are defined as

$$(4.14) \qquad \hat{h}_i = h_{2i-1} + h_{2i}, \qquad \hat{h}_{i+1} = h_{2i+1} + h_{2i+2},$$

and c_1 is a constant independent of σ_t, $\hat{h}_i$ and $\hat{h}_{i+1}$.

Proof: When $\max_{k=(2i-1,\ldots,2i+2)}(\sigma_t h_k) \ll 1$, then $\max_{k=(i,i+1)}(\sigma_t \hat{h}_k) \ll 1$. From (3.22), we can expand $\begin{bmatrix} A_i^{2h} & -C_i^{2h} \\ -D_{i+1}^{2h} & A_{i+1}^{2h} \end{bmatrix}^{-1}$ as

$$(4.15) \qquad \begin{bmatrix} A_i^{2h} & -C_i^{2h} \\ -D_{i+1}^{2h} & A_{i+1}^{2h} \end{bmatrix}^{-1} =$$

$$\begin{bmatrix}
0 & 0 & \sigma_t \hat{h}_i M^{-1} & 0 & \sigma_t \hat{h}_{i+1} M^{-1} & 0 & \sigma_t \hat{h}_{i+1} M^{-1} & 0 \\
0 & \sigma_t \hat{h}_i M^{-1} & 0 & -\sigma_t \hat{h}_i \frac{M^{-1}}{2} & 0 & 0 & 0 & 0 \\
-\sigma_t \hat{h}_i \frac{M^{-1}}{2} & 0 & \sigma_t \hat{h}_i M^{-1} & 0 & 0 & 0 & \sigma_t \hat{h}_{i+1} M^{-1} & 0 \\
0 & \sigma_t \hat{h}_i M^{-1} & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & \sigma_t \hat{h}_{i+1} M^{-1} & 0 \\
0 & \sigma_t \hat{h}_i M^{-1} & 0 & 0 & 0 & \sigma_t \hat{h}_{i+1} M^{-1} & 0 & -\sigma_t \hat{h}_{i+1} \frac{M^{-1}}{2} \\
0 & 0 & 0 & 0 & -\sigma_t \hat{h}_{i+1} \frac{M^{-1}}{2} & 0 & \sigma_t \hat{h}_{i+1} M^{-1} & 0 \\
0 & \sigma_t \hat{h}_i M^{-1} & 0 & \sigma_t \hat{h}_i M^{-1} & 0 & \sigma_t \hat{h}_{i+1} M^{-1} & 0 & 0
\end{bmatrix}$$

$$+O(\max(\sigma_t^2 \hat{h}_i^2, \sigma_t^2 \hat{h}_{i+1}^2)).$$

where M is defined by$(3.20c)$. Since M is diagonal and constant, we can find constants c_1 such that

$$(4.16) \qquad || \begin{bmatrix} A_i^{2h} & -C_i^{2h} \\ -D_{i+1}^{2h} & A_{i+1}^{2h} \end{bmatrix}^{-1} ||_{1,\infty} < c_1 \max(\sigma_t \hat{h}_i, \sigma_t \hat{h}_{i+1}).$$

$\square$

To simplify our notation, let us define

(4.17a)

$$H_i = \begin{bmatrix} A_i^{2h} & -C_i^{2h} \\ -D_{i+1}^{2h} & A_{i+1}^{2h} \end{bmatrix}^{-1} \begin{bmatrix} S_{2i-1,1} & S_{2i-1,2} & 0 & 0 \\ 0 & 0 & S_{2i+1,1} & S_{2i+1,2} \end{bmatrix} \begin{bmatrix} A_{2i-1}^h & -C_{2i-1}^h & 0 & 0 \\ -D_{2i}^h & A_{2i}^h & -C_{2i}^h & 0 \\ 0 & -D_{2i+1}^h & A_{2i+1}^h & -C_{2i+1}^h \\ 0 & 0 & -D_{2i+2}^h & A_{2i+2}^h \end{bmatrix},$$

(4.17b)

$$U_{i+1} = \begin{bmatrix} A_i^{2h} & -C_i^{2h} \\ -D_{i+1}^{2h} & A_{i+1}^{2h} \end{bmatrix}^{-1} \begin{bmatrix} S_{2i-1,1} & S_{2i-1,2} & 0 & 0 \\ 0 & 0 & S_{2i+1,1} & S_{2i+1,2} \end{bmatrix} \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ C_{2i+2} & 0 & 0 & 0 \end{bmatrix},$$

(4.17c)

$$V_{i+1} = \begin{bmatrix} A_{i+2}^{2h} & -C_{i+2}^{2h} \\ -D_{i+3}^{2h} & A_{i+3}^{2h} \end{bmatrix}^{-1} \begin{bmatrix} S_{2i+3,1} & S_{2i+3,2} & 0 & 0 \\ 0 & 0 & S_{2i+5,1} & S_{2i+5,2} \end{bmatrix} \begin{bmatrix} 0 & 0 & 0 & D_{2i+3} \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix},$$

where $A_k^{2h}, C_k^{2h}, D_k^{2h}$ are defined by (4.7), $S_{k,1}, S_{k,2}$ are defined by (4.5) and A_k^h, C_k^h, D_k^h are defined by (4.2).

Lemma 3 *When* $\max(\sigma_t h_{2i-1}, \sigma_t h_{2i}, \sigma_t h_{2i+1}, \sigma_t h_{2i+2}) \ll 1$*, then,*

$$||H_i||_{1,\infty} \le c_2 \frac{\max(\sigma_t(h_{2i-1}+h_{2i}), \sigma_t(h_{2i+1}+h_{2i+2}))}{\min(\sigma_t h_{2i-1}, \sigma_t h_{2i}, \sigma_t h_{2i+1}, \sigma_t h_{2i+2})}, \tag{4.18a}$$

$$||U_{i+1}||_{1,\infty} \le c_3 \frac{\max(\sigma_t(h_{2i-1}+h_{2i}), \sigma_t(h_{2i+1}+h_{2i+2}))}{\sigma_t(h_{2i+1}+h_{2i+2})}, \tag{4.18b}$$

$$||V_{i+1}||_{1,\infty} \le c_3 \frac{\max(\sigma_t(h_{2i+3}+h_{2i+4}), \sigma_t(h_{2i+5}+h_{2i+6}))}{\sigma_t(h_{2i+3}+h_{2i+4})}, \tag{4.18c}$$

where c_2, and c_3 are constants independent of mesh sizes.

Proof: When $\max(\sigma_t h_{2i-1}, \sigma_t h_{2i}, \sigma_t h_{2i+1}, \sigma_t h_{2i+2}) \ll 1$, $||B_k^h|| \ll 1, k = 2i-1, 2i, 2i+1, 2i+2$ in (4.2). By taking the dominant terms, we have

$$\begin{bmatrix} S_{2i-1,1} & S_{2i-1,2} & 0 & 0 \\ 0 & 0 & S_{2i,1} & S_{2i,2} \end{bmatrix} \begin{bmatrix} A_{2i-1}^h & -C_{2i-1}^h & 0 & 0 \\ -D_{2i}^h & A_{2i}^h & -C_{2i}^h & 0 \\ 0 & -D_{2i+1}^h & A_{2i+1}^h & -C_{2i+1}^h \\ 0 & 0 & -D_{2i+2}^h & A_{2i+2}^h \end{bmatrix} \tag{4.19}$$

$$= \begin{bmatrix} H_{11} & H_{12} & H_{13} & 0 \\ 0 & H_{22} & H_{23} & H_{24} \end{bmatrix} + O(1),$$

where

$$H_{11} = \begin{bmatrix} \frac{2}{\sigma_t h_{2i-1}} M & 0 & -\frac{2}{\sigma_t h_{2i-1}} M & 0 \\ 0 & 0 & 0 & 0 \\ \frac{1}{\sigma_t(h_{2i-1}+h_{2i})} M & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \tag{4.20a}$$

$$H_{12} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{\sigma_t(h_{2i-1}+h_{2i})} M \\ 0 & 0 & 0 & 0 \\ 0 & -\frac{2}{\sigma_t h_{2i}} M & 0 & \frac{2}{\sigma_t h_{2i}} M \end{bmatrix}, \tag{4.20b}$$

(4.20c)
$$H_{13} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ -\frac{1}{\sigma_t(h_{2i-1}+h_{2i})}M & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix},$$

(4.20d)
$$H_{22} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -\frac{1}{\sigma_t(h_{2i+1}+h_{2i+2})}M \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix},$$

(4.20e)
$$H_{23} = \begin{bmatrix} \frac{2}{\sigma_t h_{2i+1}}M & 0 & -\frac{2}{\sigma_t h_{2i+1}}M & 0 \\ 0 & 0 & 0 & 0 \\ \frac{1}{\sigma_t(h_{2i+1}+h_{2i+2})}M & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix},$$

(4.20f)
$$H_{24} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{\sigma_t(h_{2i+1}+h_{2i+2})}M \\ 0 & 0 & 0 & 0 \\ 0 & -\frac{2}{\sigma_t h_{2i+2}}M & 0 & \frac{2}{\sigma_t h_{2i+2}}M \end{bmatrix}.$$

Since M is constant, we can find a c_4 such that

(4.21)
$$||\begin{bmatrix} S_{2i-1,1} & S_{2i-1,2} & 0 & 0 \\ 0 & 0 & S_{2i,1} & S_{2i,2} \end{bmatrix} \begin{bmatrix} A^h_{2i-1} & -C^h_{2i-1} & 0 & 0 \\ -D^h_{2i} & A^h_{2i} & -C^h_{2i} & 0 \\ 0 & -D^h_{2i+1} & A^h_{2i+1} & -C^h_{2i+1} \\ 0 & 0 & -D^h_{2i+2} & A^h_{2i+2} \end{bmatrix}||_{1,\infty}$$
$$\leq \frac{c_4}{\min(\sigma_t h_{2i-1}, \sigma_t h_{2i}, \sigma_t h_{2i+1}, \sigma_t h_{2i+2})}.$$

Combined with *Lemma* 2, we have

(4.22)
$$||H_i||_{1,\infty} \leq c_2 \frac{\max(\sigma_t(h_{2i-1}+h_{2i}), \sigma_t(h_{2i+1}+h_{2i+2}))}{\min(\sigma_t h_{2i-1}, \sigma_t h_{2i}, \sigma_t h_{2i+1}, \sigma_t h_{2i+2})},$$

with $c_2 = c_1 c_4$.

To prove $(4.18b)$, we have

(4.23)
$$\begin{bmatrix} S_{2i-1,1} & S_{2i-1,2} & 0 & 0 \\ 0 & 0 & S_{2i+1,1} & S_{2i+1,2} \end{bmatrix} \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ C_{2i+2} & 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ U_{21} & 0 & 0 & 0 \end{bmatrix},$$

with

$$U_{21} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ \frac{1}{\sigma_t(h_{2i+1}+h_{2i+2})}M & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}. \tag{4.24}$$

Using *Lemma* 2, (4.18b) is proved. The final result, (4.18c), can be proved in the same manner. □

We next assume that on the finest grid, the value of $\sigma_t h_i$ does not vary too quickly; that is, we assume that for each i

$$\frac{\max(\sigma_t(h_{2i-1}+h_{2i}), \sigma_t(h_{2i+1}+h_{2i+2}))}{\min(\sigma_t h_{2i-1}, \sigma_t h_{2i}, \sigma_t h_{2i+1}, \sigma_t h_{2i+2})} \le c_5, \tag{4.25}$$

where c_5 is a bound independent of mesh sizes. With (4.25), we have

$$||H_i||_{1,\infty} \le c_6, \quad ||U_{i+1}||_{1,\infty} \le c_7, \quad ||V_{i+1}||_{1,\infty} \le c_7, \tag{4.26}$$

where $c_6 = c_2 c_5$ and $c_7 = c_3 c_5$.

We arrange (4.1) as

$$A^{2h} = Z_1 - Z_2, \tag{4.27}$$

where

$$Z_1 = \begin{bmatrix} A_1^{2h} & -C_1^{2h} & & & & & & \\ -D_2^{2h} & A_2^{2h} & 0 & & & & & \\ & 0 & \ddots & & & & & \\ & & & A_i^{2h} & -C_i^{2h} & & & \\ & & & -D_{i+1}^{2h} & A_{i+1}^{2h} & & & \\ & & & & & \ddots & 0 & \\ & & & & & 0 & A_{\frac{m}{2}-1}^{2h} & -C_{\frac{m}{2}-1}^{2h} \\ & & & & & & -D_{\frac{m}{2}}^{2h} & A_{\frac{m}{2}}^{2h} \end{bmatrix}, \tag{4.28a}$$

$$Z_2 = \begin{bmatrix} 0 & 0 & 0 & 0 & & & & \\ 0 & 0 & C_2^{2h} & 0 & & & & \\ 0 & D_3^{2h} & 0 & 0 & \ddots & & & \\ 0 & 0 & & \ddots & & \ddots & & \\ & & \ddots & & 0 & 0 & 0 & 0 \\ & & & \ddots & 0 & 0 & C_{\frac{m}{2}-2}^{2h} & 0 \\ & & & & 0 & D_{\frac{m}{2}-1}^{2h} & 0 & 0 \\ & & & & 0 & 0 & 0 & 0 \end{bmatrix}. \tag{4.28b}$$

Then

$$(A^{2h})^{-1} = (I - Z_1^{-1} Z_2)^{-1} Z_1^{-1}. \tag{4.29}$$

Lemma 4 *Suppose* $\max\limits_{k=(1,\ldots,\frac{m}{2})} (\sigma_t \hat{h}_k) \ll 1$ *and* (4.25) *holds for all* i*; then* $||(I - Z_1^{-1} Z_2)^{-1}||_2$ *is bounded by a constant independent of* $\sigma_t h_i$.

Proof: By using (4.15), the leading term of $I - Z_1^{-1} Z_2$ can be expanded as

$$\begin{bmatrix} I & Q_1 & & \\ Q_2 & I & \ddots & \\ & \ddots & \ddots & Q_1 \\ & & Q_2 & I \end{bmatrix}, \tag{4.30}$$

where

$$Q_1 = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \tag{4.31a}$$

$$Q_2 = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}. \tag{4.31b}$$

Notice that (4.30) can be obtained by performing elementary operations on an identity matrix. So

$$(I - Z_1^{-1} Z_2) = P_1 I P_2 + O(\max_{k=(1,\ldots,\frac{m}{2})} (\sigma_t \hat{h}_k)), \tag{4.32}$$

where P_1 and P_2 are elementary matrices independent of $\sigma_t \hat{h}_i$. Thus

$$||(I - Z_1^{-1} Z_2)^{-1}||_2 < c_8, \tag{4.33}$$

where c_8 is a constant independent of $\sigma_t \hat{h}_i$. $\square$

With all the lemmas established, we present in the next Theorem an upper bound on the convergence factor of multigrid algorithm when $\max_{k=(1,\ldots,m)} (\sigma_t h_k) \ll 1$.

Theorem 5 *When* $\max_{k=(1,\ldots,m)} (\sigma_t h_k) \ll 1$ *and* (4.25) *holds for all* i*, a two-grid multigrid* $V(1,0)$ *with two-cell red-black block* μ*-line relaxation will have a convergence factor which is less than* $O(\max_{k=(1,\ldots,m)} (\sigma_t^2 h_k^2))$.

Proof: From Theorem 2, the errors after two-cell red-black block μ-line relaxation can be expressed as

$$\underline{e}_1^h = I_{2h}^h \underline{e}^{2h} + O(\max_{k=(1,\ldots,m)} (\sigma_t^2 h_k^2))\underline{e}_0^h. \tag{4.34}$$

A multigrid $V(1,0)$ cycle in matrix form takes the form of

$$\underline{e}_2^h = G^h \underline{e}_0^h - I_{2h}^h (A^{2h})^{-1} I_h^{2h} A^h G^h \underline{e}_0^h. \tag{4.35}$$

Substitute (4.34) into (4.35), we have

$$\begin{aligned} \underline{e}_2^h &= I_{2h}^h \underline{e}^{2h} + O(\max_{k=(1,\ldots,m)} (\sigma_t^2 h_k^2))\underline{e}_0^h - I_{2h}^h (A^{2h})^{-1} I_h^{2h} A^h I_{2h}^h \underline{e}^{2h} \\ &\quad + I_{2h}^h (A^{2h})^{-1} I_h^{2h} A^h O(\max_{k=(1,\ldots,m)} (\sigma_t^2 h_k^2))\underline{e}_0^h \\ &= O(\max_{k=(1,\ldots,m)} (\sigma_t^2 h_k^2))\underline{e}_0^h + I_{2h}^h (A^{2h})^{-1} I_h^{2h} A^h O(\max_{k=(1,\ldots,m)} (\sigma_t^2 h_k^2))\underline{e}_0^h. \end{aligned} \tag{4.36}$$

Note that

$$I_{2h}^h \underline{e}^{2h} = I_{2h}^h (A^{2h})^{-1} I_h^{2h} A^h I_{2h}^h \underline{e}^{2h}. \tag{4.37}$$

It is obvious that we only need to bound the norm of $I_{2h}^h (A^{2h})^{-1} I_h^{2h} A^h$.

From (4.29), we obtain

$$(A^{2h})^{-1} I_h^{2h} A^h = (I - Z_1^{-1} Z_2)^{-1} Z_1^{-1} I_h^{2h} A^h, \tag{4.38}$$

and

$$Z_1^{-1} I_h^{2h} A^h = \begin{bmatrix} H_1 & -U_2 & & & \\ -V_2 & H_3 & & & \\ & & \ddots & & \\ & & & -V_{m-2} & H_{m-1} \end{bmatrix}. \tag{4.39}$$

By our assumption of (4.25) and (4.26), we have

$$||Z_1^{-1} I_h^{2h} A^h||_{1,\infty} \le c_9. \tag{4.40}$$

For any matrix A, $||A||_2 \le \sqrt{||A||_1 ||A||_\infty}$. So

$$||Z_1^{-1} I_h^{2h} A^h||_2 \le c_9. \tag{4.41}$$

The norms of I_{2h}^h are bounded by a constant. So

$$||I_{2h}^h||_2 < c_{10}. \tag{4.42}$$

Combining *Lemma* 4 with Equations (4.41) and (4.42), we obtain

$$||\underline{e}_2^h||_2 / ||\underline{e}_0^h||_2 < O(\max_{k=(1,\ldots,m)} (\sigma_t^2 h_k^2)). \tag{4.43}$$

□

When $\min_{k=(1,\ldots,m)} (\sigma_t h_k) \gg 1$, we limit our analysis to only four cells on the fine grid with uniform mesh size h and our numerical results imply that the analytical result is valid with nonuniform mesh size and large number of cells. The next theorem gives an upper bound on convergence factor of multigrid algorithm when $m = 4$ and $\sigma_t h \gg 1$. First, let us establish some space saving notation. Define the block matrices

$$E_{j,l} = [0, ..., 0, I, 0, ..., 0], \tag{4.44}$$

to have l blocks of size $n \times n$ with the I appearing as the jth block. Then we define another block matrix with $n \times n$ blocks, say

$$W = [W_1, ..., W_k]. \tag{4.45}$$

The product $W^T E_{j,l}$ is a $k \times l$ block matrix with $n \times n$ blocks. Each block column is zero except the jth block column which contains W^T. For example, we may write C_i in $(2.18b)$ as

$$C_i = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ B_i & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} = [0, 0, B_i, 0]^T E_{1,4}. \tag{4.46}$$

Theorem 6: When there are four uniform cells on the fine grid and $\sigma_t h \gg 1$, a multigrid $V(1,1)$ cycle with two-cell red-black block μ-line relaxation will have a convergence factor $\rho \le c(\frac{1}{\sigma_t h})$ for some c independent of $\sigma_t h$.

Proof: Define A^h as the fine grid operator, then

$$A^h = \begin{bmatrix} A & -C \\ -D & A \end{bmatrix}, \tag{4.47}$$

where

(4.48a)

$$A = \left[\begin{array}{cccccccc} I+2B-R & -2R & -2B & R & 0 & 0 & 0 & 0 \\ 0 & I-R & -R & B & 0 & 0 & 0 & 0 \\ B & -R & I-R & 0 & -B & 0 & 0 & 0 \\ R & -2B & -2R & I+2B-R & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & I+2B-R & -2R & -2B & R \\ 0 & 0 & 0 & -B & 0 & I-R & -R & B \\ 0 & 0 & 0 & 0 & B & -R & I-R & 0 \\ 0 & 0 & 0 & 0 & R & -2B & -2R & I+2B-R \end{array} \right],$$

and

(4.48b)

$$C = \left[\begin{array}{cccccccc} 0 & 0 & 0 & 0 & 0 & 0 & B & 0 \end{array} \right]^T E_{1,8}, \quad D = \left[\begin{array}{cccccccc} 0 & B & 0 & 0 & 0 & 0 & 0 & 0 \end{array} \right]^T E_{8,8}.$$

The coarse grid operator A^{2h} is defined by

$$A^{2h} = I_h^{2h} A^h I_{2h}^h, \tag{4.49}$$

with I_h^{2h} and I_{2h}^h defined by

(4.50a)

$$I_h^{2h} = \left[\begin{array}{cccc} S_1 & S_2 & 0 & 0 \\ 0 & 0 & S_1 & S_2 \end{array} \right], S_1 = \left[\begin{array}{cccc} I & 0 & 0 & 0 \\ 0 & 0.5I & 0 & 0 \\ 0 & 0 & 0.5I & 0 \\ 0 & 0 & 0 & 0 \end{array} \right], S_2 = \left[\begin{array}{cccc} 0 & 0 & 0 & 0 \\ 0 & 0.5I & 0 & 0 \\ 0 & 0 & 0.5I & 0 \\ 0 & 0 & 0 & I \end{array} \right],$$

(4.50b)

$$I_{2h}^h = \left[\begin{array}{cc} T_1 & 0 \\ T_2 & 0 \\ 0 & T_1 \\ 0 & T_2 \end{array} \right], T_1 = \left[\begin{array}{cccc} I & 0 & 0 & 0 \\ 0 & 1.5I & 0 & -0.5I \\ 0.5I & 0 & 0.5I & 0 \\ 0 & I & 0 & 0 \end{array} \right], T_2 = \left[\begin{array}{cccc} 0 & 0 & I & 0 \\ 0 & 0.5I & 0 & 0.5I \\ -0.5I & 0 & 1.5I & 0 \\ 0 & 0 & 0 & I \end{array} \right].$$

The red-black relaxation matrix G^h is

$$G^h = \left[\begin{array}{cc} A & 0 \\ -D & A \end{array} \right]^{-1} \left[\begin{array}{cc} 0 & C \\ 0 & 0 \end{array} \right] = \left[\begin{array}{cc} 0 & A^{-1}C \\ 0 & A^{-1}DA^{-1}C \end{array} \right]. \tag{4.51}$$

A $V(1,1)$ multigrid cycle can be written in matrix form as

$$\underline{e}_2^h = G^h(G^h - I_{2h}^h (A^{2h})^{-1} I_h^{2h} A^h G^h) \underline{e}_0^h, \tag{4.52}$$

here $\underline{e}_0^h$ is the initial error and $\underline{e}_2^h$ is the error after the $V(1,1)$ cycle.

Multiply G^h by A^h on the left to get

$$A^h G^h = \left[\begin{array}{cc} 0 & C - CA^{-1}DA^{-1}C \\ 0 & 0 \end{array} \right] = \left[\begin{array}{cc} C & 0 \\ 0 & 0 \end{array} \right] \left[\begin{array}{cc} 0 & I - A^{-1}DA^{-1}C \\ 0 & 0 \end{array} \right]. \tag{4.53}$$

Note that

$$I_h^{2h}\begin{bmatrix} C & 0 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 & \frac{1}{2}B & 0 & 0 & 0 & 0 & 0 \end{bmatrix}^T E_{1,16}. \tag{4.54}$$

The expansion of $(A^{2h})^{-1}I_h^{2h}\begin{bmatrix} C & 0 \\ 0 & 0 \end{bmatrix}$ is

$$(A^{2h})^{-1}I_h^{2h}\begin{bmatrix} C & 0 \\ 0 & 0 \end{bmatrix} = \frac{\sigma_t h}{4c_2}\begin{bmatrix} 0 \\ RM \\ RM \\ 2RM \\ 2RM \\ RM \\ RM \\ RM \\ 0 \end{bmatrix} E_{1,16} \tag{4.55}$$

$$+\frac{1}{2c_1}\begin{bmatrix} \frac{1}{2}RM + \frac{c_1}{2c_2}MRM \\ \frac{1}{8}(3+\frac{3c_1c_3}{c_2^2})RM - \frac{c_1}{2c_2}MRM - \frac{c_1}{4c_2}RM^2 \\ \frac{1}{8}(3+\frac{3c_1c_3}{c_2^2})RM + \frac{c_1}{2c_2}MRM - \frac{c_1}{4c_2}RM^2 \\ \frac{1}{4}(3+\frac{3c_1c_3}{c_2^2})RM - \frac{c_1}{2c_2}MRM - \frac{c_1}{2c_2}RM^2 \\ \frac{1}{4}(-1+\frac{3c_1c_3}{c_2^2})RM - \frac{c_1}{2c_2}MRM - \frac{c_1}{2c_2}RM^2 \\ \frac{1}{8}(1+\frac{3c_1c_3}{c_2^2})RM + \frac{c_1}{2c_2}MRM - \frac{c_1}{4c_2}RM^2 \\ \frac{1}{8}(1+\frac{3c_1c_3}{c_2^2})RM - \frac{c_1}{4c_2}MRM - \frac{c_1}{4c_2}RM^2 \\ \frac{1}{2}RM + \frac{c_1}{2c_2}MRM \end{bmatrix} E_{1,16} + O(\frac{1}{\sigma_t h}),$$

where $M = \sigma_t hB$ and $c_1 = \sum_{j=1}^n \omega_j\mu_j$, $c_2 = \sum_{j=1}^n \omega_j\mu_j^2$ and $c_3 = \sum_{j=1}^n \omega_j\mu_j^3$. Note $RMR = c_1R$, $RM^2R = c_2R$ and $RM^3R = c_3R$.

Expand $I - A^{-1}DA^{-1}C$ as

$$I - A^{-1}DA^{-1}C = I - \frac{1}{4c_1}\begin{bmatrix} 4RM \\ 3RM \\ 3RM \\ 2RM \\ 2RM \\ RM \\ RM \\ 0 \end{bmatrix} E_{1,8}. \tag{4.56}$$

$$-\frac{1}{8c_1^2\sigma_t h}\begin{bmatrix} -8c_2RM-4c_1MRM-4c_1RM^2 \\ -5c_2RM+4c_1MRM-\frac{1}{4}(13-\frac{c_1c_3}{c_2^2})c_1RM^2 \\ -5c_2RM-4c_1MRM-\frac{1}{4}(13-\frac{c_1c_3}{c_2^2})c_1RM^2 \\ -2c_2RM+4c_1MRM-\frac{1}{2}(5-\frac{c_1c_3}{c_2^2})c_1RM^2 \\ -2c_2RM-4c_1MRM-\frac{1}{2}(5-\frac{c_1c_3}{c_2^2})c_1RM^2 \\ c_2RM-4c_1MRM-\frac{1}{4}(5-\frac{c_1c_3}{c_2^2})c_1RM^2 \\ c_2RM+4c_1MRM-\frac{1}{4}(5-\frac{c_1c_3}{c_2^2})c_1RM^2 \\ 4c_2RM+4c_1MRM \end{bmatrix} E_{1,8}+O(\frac{1}{\sigma_t^2h^2})$$

From (4.55) and (4.56), we get

$$(A^{2h})^{-1}I_h^{2h}\begin{bmatrix} C & 0 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} 0 & I-A^{-1}DA^{-1}C \\ 0 & 0 \end{bmatrix}=\frac{1}{4c_1}\begin{bmatrix} 0 \\ 2RM \\ 2RM \\ 4RM \\ 4RM \\ 2RM \\ 2RM \\ 0 \end{bmatrix}E_{9,16}+O(\frac{1}{\sigma_t h}). \tag{4.57}$$

Then

$$I_{2h}^h(A^{2h})^{-1}I_h^{2h}A^hG^h=\frac{1}{4c_1}\begin{bmatrix} 0 \\ RM \\ RM \\ 2RM \\ 2RM \\ 3RM \\ 3RM \\ 4RM \\ 4RM \\ 3RM \\ 3RM \\ 2RM \\ 2RM \\ RM \\ RM \\ 0 \end{bmatrix}E_{9,16}+O(\frac{1}{\sigma_t h}). \tag{4.58}$$

The expansion of G^h takes the form of

$$G^h = \frac{1}{4c_1}\begin{bmatrix} 0 \\ \frac{1}{2}RM + \frac{c_1}{2c_2}RM^2 \\ \frac{1}{2}RM + \frac{c_1}{2c_2}RM^2 \\ RM + \frac{c_1}{c_2}RM^2 \\ RM + \frac{c_1}{c_2}RM^2 \\ \frac{5}{2}RM + \frac{c_1}{2c_2}RM^2 \\ \frac{5}{2}RM + \frac{c_1}{2c_2}RM^2 \\ 4RM \\ 4RM \\ 3RM \\ 3RM \\ 2RM \\ 2RM \\ RM \\ RM \\ 0 \end{bmatrix} E_{9,16} + O(\frac{1}{\sigma_t h}). \tag{4.59}$$

So

$$G^h G^h = \frac{1}{4c_1}\begin{bmatrix} 0 \\ RM \\ RM \\ 2RM \\ 2RM \\ 3RM \\ 3RM \\ 4RM \\ 4RM \\ 3RM \\ 3RM \\ 2RM \\ 2RM \\ RM \\ RM \\ 0 \end{bmatrix} E_{9,16} + O(\frac{1}{\sigma_t h}) \tag{4.60}$$

and

$$
G^h I_{2h}^h (A^{2h})^{-1} I_h^{2h} A^h G^h = \frac{1}{4c_1} \begin{bmatrix} 0 \\ RM \\ RM \\ 2RM \\ 2RM \\ 3RM \\ 3RM \\ 4RM \\ 4RM \\ 3RM \\ 3RM \\ 2RM \\ 2RM \\ RM \\ RM \\ 0 \end{bmatrix} E_{9,16} + O(\frac{1}{\sigma_t h}). \tag{4.61}
$$

Combine (4.60) and (4.61), we can have

$$
||G^h(G^h - I_{2h}^h (A^{2h})^{-1} I_h^{2h} A^h G^h)|| \leq \frac{c}{\sigma_t h}. \tag{4.62}
$$

here c is a constant independent of $\sigma_t h$. $\square$

§5 **Computational Results**

5.1 Multigrid Convergence Rates

The following computational results were conducted on a computational domain with $m = 1024$ and S_8, that is $n = 4$. Our analytical results are independent of n. From the computational results conducted on S_N with N ranging from 2 to 256, we observe a convergence performance independent of N.

1. Convergence Factor for Uniform Grid

The results in Table 1 reveal that the convergence factor ρ is $O(\sigma_t^3 h^3)$ when $\sigma_t h \ll 1$ and is $O(\frac{1}{\sigma_t^2 h^2})$ when $\sigma_t h \gg 1$. The computational results are even better than the analytical results. Our analysis only gives upper bounds on the convergence factor. We believe that the numerical results reflect the actual behavior. In Table 2, we present a range within which the maximal convergence factor occurs. Note that the maximal ρ occurs around $\sigma_t h = 0.01$.

σ_t h	convergence factor ρ
10^{-6}	0.23×10^{-10}
10^{-5}	0.21×10^{-7}
10^{-4}	0.16×10^{-4}
10^{-3}	0.20×10^{-2}
10^{-2}	0.97×10^{-2}
10^{-1}	0.89×10^{-2}
1	0.58×10^{-3}
10^{1}	0.27×10^{-4}
10^{2}	0.20×10^{-6}
10^{3}	0.20×10^{-8}
10^{4}	0.20×10^{-10}

Table 1. Convergence Factors for Uniform Grid

σ_t h	convergence factor ρ
0.005	0.72×10^{-2}
0.006	0.87×10^{-2}
0.007	0.95×10^{-2}
0.008	0.98×10^{-2}
0.009	0.98×10^{-2}
0.010	0.97×10^{-2}
0.011	0.96×10^{-2}
0.012	0.96×10^{-2}
0.013	0.95×10^{-2}
0.014	0.95×10^{-2}
0.015	0.95×10^{-2}

Table 2. Convergence Factors for Uniform Grid: Worst Case

2. Convergence Factor for Nonuniform Grid

We choose a highly variable $\sigma_t h_i$ by letting $\sigma_t h_i = c10^{2\eta_i}, i = 1, ..., m$ where η_i is a random number between $(-1, 1)$. Thus, neighboring cells can vary in width up to 4 orders of magnitude. Here, c is a constant chosen to control the range of cell widths. If $c = 1$, then $\sigma_t h_i$ are random numbers between (0.01,100). If $c \geq 100$, $\sigma_t h_i \geq 1$. If $c \leq 0.01$, all $\sigma_t h_i \leq 1$. The results of Table 3 indicate that the performance of our multigrid scheme for a nonuniform grid is the same as for a uniform grid. Furthermore in both thin and thick limit cases, a nonuniform grid has the same convergence behavior as uniform grid.

c	convergence factor ρ
1.0^{-6}	9.9×10^{-10}
1.0^{-5}	8.5×10^{-7}
1.0^{-4}	2.7×10^{-4}
1.0^{-3}	5.4×10^{-3}
1.0	9.1×10^{-4}
1.0^{2}	3.3×10^{-4}
1.0^{3}	6.1×10^{-6}
1.0^{4}	2.4×10^{-8}
1.0^{5}	2.3×10^{-10}

Table 3. Convergence Factors for Nonuniform Grids

5.2 DSA Convergence Rates

As was mentioned in the introduction, a competing algorithm for the solution of the isotropic S_N equations in slab geometry is Diffusion Synthetic Acceleration (DSA). It is motivated by the observation that, in the thick limit, the solution of the linear Boltzman equation (1.1) becomes independent of angle except near boundaries and sources. If the S_N problem were recast as an S_2 problem, then the zeroth and first Legendre moments of the solution could be found by solving a diffusion equation for the zeroth moment. DSA can be viewed as a preconditioning technique (Faber and Manteuffel [5]) or as a two level multigrid scheme (Larson [7]) in which an S_2 problem is used as a coarse level correction to an S_N problem. In this context, a transport sweep or x-line relaxation (see(3.3)) corresponds to multigrid relaxation and the S_2 problem represents the coarse level. DSA skips all of the intermediate levels. The multigrid in angle scheme developed in Morel and Manteuffel [15] visits levels $N, N/2, N/4, \ldots, 2$ and is effective for isotropic scattering as well as highly anisotropic scattering.

The overall DSA algorithm is very sensitive to the difference scheme used to solve the diffusion equation. In our numerical results we solve an S_2 equation discretized with a MLD scheme using Marshak boundary conditions as the coarse level correction. This is equivalent to solving a consistently differenced diffusion equation for the zeroth moment. In Table 4 and Table 5 we present convergence factors for a $V(1,1)$ multigrid cycle and for a DSA cycle. Here, the slab is assumed to have physical thickness 1. Thus, σ_t represents the width of the slab measured in the number of mean-free-paths. The tests were performed using S_8 and a wide range of σ_t and m, (the number of cells). The diagonals of these tables represent constant $\sigma_t h$.

σ_t	$m = 16$	$m = 64$	$m = 256$	$m = 1024$	m = 4096
4^{-5}	.12E-9	.16E-9	.16E-9	.16E-9	.16E-9
4^{-4}	.72E-8	.10E-7	.10E-7	.10E-7	.10E-7
4^{-3}	.37E-6	.55E-6	.56E-6	.56E-6	.56E-6
4^{-2}	.17E-4	.16E-4	.18E-4	.26E-4	.27E-4
4^{-1}	.25E-3	.53E-3	.55E-3	.55E-4	.58E-4
1	.14E-2	.37E-2	.44E-2	.44E-2	.46E-2
4	.82E-3	.35E-2	.82E-2	.99E-2	.99E-2
16	.57E-3	.13E-2	.63E-2	.86E-2	.97E-2
64	.46E-4	.57E-3	.23E-2	.85E-2	.99E-2
256	.20E-5	.46E-4	.58E-3	.25E-2	.89E-2
1024	.12E-6	.18E-5	.47E-4	.58E-3	.26E-2
4096	.75E-8	.12E-6	.16E-5	.46E-4	.58E-3
8192	.45E-9	.75E-8	.12E-6	.14E-5	.47E-4
16384	.28E-10	.46E-9	.75E-8	.12E-6	.11E-5
32768	.18E-11	.30E-10	.47E-9	.75E-8	.12E-6
65536	.11E-12	.20E-11	.30E-10	.46E-9	.75E-8

Table 4. Convergence Factors for the Multigrid Algorithm

σ_t	$m = 16$	$m = 64$	$m = 256$	$m = 1024$	$m = 4096$
4^{-5}	.190E-2	.190E-2	.190E-2	.190E-2	.190E-2
4^{-4}	.747E-2	.746E-2	.746E-2	.747E-2	.747E-2
4^{-3}	.276E-1	.276E-1	.276E-1	.276E-1	.276E-1
4^{-2}	.836E-1	.834E-1	.834E-1	.835E-1	.835E-1
4^{-1}	.153	.151	.151	.152	.152
1	.188	.203	.204	.204	.205
4	.180	.217	.216	.217	.218
16	.142	.164	.204	.217	.219
64	.129	.143	.149	.165	.185
256	.125	.131	.143	.145	.160
1024	.124	.126	.131	.143	.146
4096	.124	.124	.126	.131	.142
8192	.124	.124	.126	.127	.132
16384	.124	.124	.124	.126	.126
32768	.124	.124	.124	.124	.126
65536	.124	.124	.124	.124	.124

Table 5. Convergence Factors for the DSA Algorithm

For the multigrid algorithm, the convergence factors are roughly equal along diagonals, with the slowest rates occurring for $\sigma_t h$ in the range $4^{-3}\,(.0156)$ to $4^{-5}\,(.97E-3)$. For example, $\sigma_t = 4^{-1}$, $m = 16$ and $\sigma_t = 64$, $m = 4096$ are on the diagonal with $\sigma_t h = .0156$.

The behavior of the DSA algorithm is quite different. The best rate achieved for thick problems is .124. Further, for any fixed $\sigma_t \geq 4$, the slab is thick but the cells become individually thin as the number of cells increases. In this limit DSA saturates to a convergence factor of .23. In our tests the the value of m was not quite large enough to achieve this limit, but a value of .219 was reached for $\sigma_t = 16$ and $m = 4096$.

In all cases, the multigrid convergence factors were superior to the DSA convergence factors. However, it is important to adjust for the relative amount of computational work required by each algorithm. Of course, such measures will be machine dependent. On the Cray Y/MP, where these tests were performed, we compared times for N=64 and m=1024. The V(1,1) cycle required 8.9 seconds and the DSA cycle required 3.7 seconds. The ratio of these times is 2.4. Table 6 is the result of raising each term in Table 5 to the power 2.4. This is a more fair comaprison with Table 4 on a serial machine. Note that the multigrid algorithm is faster than DSA in all regimes. On a parallel machine we expect the results to more heavily favor the multigrid algorithm. Both algorithms can be implemented with parallel complexity $O(log(m))$. In this context, however, we expect the times required to perform a single DSA cycle and a single multigrid V(1,1) cycle to be more nearly equal. A parallel version of the multigrid algorithm is described in [10], [11]. A parallel version of the DSA sweep has also been implemented on the Thinking Machines Inc. CM-200 at Los Alamos National Laboratories [20].

σ_t	$m = 16$	$m = 64$	$m = 256$	$m = 1024$	$m = 4096$
4^{-5}	.29E-6	.29E-6	.29E-6	.29E-6	.29E-6
4^{-4}	.78E-5	.78E-5	.78E-5	.78E-5	.78E-5
4^{-3}	.18E-3	.18E-3	.18E-3	.18E-3	.18E-3
4^{-2}	.25E-2	.25E-2	.25E-2	.25E-2	.25E-2
4^{-1}	.11E-1	.10E-1	.10E-1	.10E-1	.10E-1
1	.18E-1	.21E-1	.22E-1	.22E-1	.22E-1
4	.16E-1	.25E-1	.25E-1	.25E-1	.25E-1
16	.92E-2	.13E-1	.22E-1	.25E-1	.26E-1
64	.73E-2	.93E-2	.10E-1	.13E-1	.17E-1
256	.68E-2	.76E-2	.93E-2	.97E-2	.12E-1
1024	.66E-2	.69E-2	.76E-2	.93E-2	.98E-2
4096	.66E-2	.66E-2	.69E-2	.76E-2	.92E-2
8192	.66E-2	.66E-2	.69E-2	.70E-2	.77E-2
16384	.66E-2	.66E-2	.66E-2	.69E-2	.69E-2
32768	.66E-2	.66E-2	.66E-2	.66E-2	.69E-2
65536	.66E-2	.66E-2	.66E-2	.66E-2	.66E-2

Table 6. Adjusted Convergence Factors for the DSA Algorithm

In either setting, parallel or serial, a full multigrid algorithm can be implemented (cf.[13]). This version of multigrid starts on a coarse grid and moves toward finer grids. In general, an amount of work equal to two V(1,1) cycles on the finest grid will yield a solution that is accurate to the level of discretization error. This would provide a savings in some regions of the tables. Moreover, full multigrid provides a natural framework for adaptive grid refinement.

§6 Conclusions

From our numerical and analytical results, we conclude that

- For S_2, two-cell, red-black, block μ-line relaxation is a good smoother which makes the errors exactly in the range of interpolation after one relaxation. Thus, one multigrid $V(1,0)$ cycle is an exact solver for the S_2 problem.
- For S_N with $N > 2$, the errors at two neighboring cells $2i-1$, $2i$ after one such relaxation will be in the range of interpolation up to accuracy of $O(\max(\frac{1}{\sigma_t h_{2i-1}}, \frac{1}{\sigma_t h_{2i}}))$ when $\min(\sigma_t h_{2i-1}, \sigma_t h_{2i}) \gg 1$ and $O(\max(\sigma_t^2 h_{2i-1}^2, \sigma_t^2 h_{2i}^2))$ when $\max(\sigma_t h_{2i-1}, \sigma_t h_{2i}) \ll 1$.
- The multigrid algorithm is an efficient solver which has a convergence factor of $O(\max_{k=(1,\ldots,m)}(\frac{1}{\sigma_t^2 h_k^2}))$ for the thick limit and $O(\max_{k=(1,\ldots,m)}(\sigma_t^3 h_k^3))$ for the thin limit with a maximal of $\rho = 0.0098$ occurring around $\sigma_t h = 0.01$ for the uniform grid.
- The multigrid algorithm performs equally well for uniform and nonuniform grids.
- In the thick limit a single V(1,1) yields nearly an exact solution.
- The multigrid algorithm is more efficient than DSA in all regimes.
- The relaxation is implemented in a red-black ordering and is parallelizable.

The above results assumed no absorption. The problem is easy to solve if the absorption is sufficiently large, that is if $\gamma = \frac{\sigma_s}{\sigma_t}$ is bounded away from unity. In this case, smoothing alone is sufficient for a fast solution. With no absorption, $\gamma = 1$, the algorithm presented above provides nearly an exact solver. However, for the case in which $\gamma = 1 - O(\frac{1}{(\sigma_t h)^2})$ in the thick limit, the above algorithm yields a convergence factor that does not go to zero in the thick limit. A minor adaptation to the algorithm has been developed that treats this special case and provides a convergence factor that does go to zero in the thick limit for all levels of absorption. It will appear in a subsequent report.

References

[1] M. L. Adams and W. R. Martin, "Diffusion-Synthetic Acceleration of Discontinuous Finite-Element Transport Iterations," *Nucl. Sci. Eng.*, **111**, 145-167 (1992).

[2] R. E. Alcouffe, "Diffusion Synthetic Acceleration Methods for the Diamond-Differenced Discrete-Ordinates Equations," *Nucl. Sci. Eng.*, **64**, 344 (1977).

[3] Allen Barnett, J. E. Morel, and D. R. Harris, "A Multigrid Acceleration Method for the One-Dimensional S_n Equations with Anisotropic Scattering," *Nucl. Sci. Eng.*, **102**, 1 (1989).

[4] A. Brandt, Multi-level Adaptive Solution to Boundary-value Problems.*Math. Comp.*, **31**, 333 (1977).

[5] V. Faber and T. A. Manteuffel, "Neutron Transport from the Viewpoint of Linear Algebra Integral," *Transport Theory, Invariant Imbedding and Equations*, (Nelson, Faber, Manteuffel, Seth and White, eds.), Lecture Notes in Pure and Applied Mathematics, **115**, pp. 37-61, Marcel-Dekker, April, 1989.

[6] E. W. Larsen, "Unconditionally Stable Diffusion-Synthetic Acceleration Methods for the Slab Geometry Discrete Ordinates Equations," *Nucl. Sci. Eng.*, **82**, 47 (1982).

[7] E. W. Larsen, "Transport Acceleration Methods as Two-Level Multigrid Algorithms," *Operator Theory: Advances and Applications*, Vol. 51, p34-47, (1991).

[8] E.W.Larsen, J.E. Morel, "Asymptotic Solutions of Numerical Transport Problems in Optically Thick Diffusive Regimes II," *J. Comp. Phys.*, **83**, 212 (1989).

[9] E. E. Lewis and W. F. Miller,Jr., *Computational Methods of Neutron Transport*, John Wiley and Sons, (1984).

[10] T. A. Manteuffel, S. F. McCormick, J. E. Morel, S. Oliveira, and G. Yang, "Parallel Multigrid Methods for Transport Equations," *Proceedings of the Copper Mountain Conference on Iterative Methods, Copper Mountain, CO, April 9 - 14, 1992*

[11] T. A. Manteuffel, S. F. McCormick, J. E. Morel, S. Oliveira, and G. Yang, " A Parallel Version of a Multigrid Algorithm for Isotropic Transport Equations," *SIAM J. Sci Stat. Comp.*, Vol. 15, No. 2, (1994)

[12] T. A. Manteuffel, S.F. McCormick, J. E. Morel and G. Yang, " A Fast Multigrid Algorithm for Transport Problems II: With Absorption," submitted to *SIAM J. Sci. and Stat. Comp.*,, March, 1993.

[13] S. F. McCormick, *Multilevel Adaptive Methods for Partial Differential Equations*, SIAM Press, (1989).

[14] J. E. Morel and E. W. Larsen, "A Multiple Balance Approach for Differencing the S_N Equations," *Nucl. Sci. Eng.*, **105**, 1 (1990).

[15] J. E. Morel and T. A. Manteuffel: "An Angular Multigrid Acceleration Technique for the S_N Equations With Highly Forward-Peaked Scattering," *Nucl. Sci. Eng.*, **107** 330 (1991).

[16] P. F. Nowak, " A Coupled Synthetic and Multigrid Acceleration Method for Two-Dimensional Transport Calculations PhD thesis, Department of Nuclear Engineering, University of Michigan, 1988.

[17] P. F. Nowak, *A Multigrid Method for S_N Calculations in x-y Geometry*, Transactions of the American Nuclear Society, vol. 56 pp. 291-292, 1988.

[18] P. F. Nowak, E. W. Larsen, *Multigrid methods for S_N problems*, Transactions of the American Nuclear Society, vol. 57, pp. 355-356, 1987.

[19] A. Nowak, J. E. Morel and D.R.Harris, *A Multigrid Acceleration Method for one dimensional S_N Equations with Anisotropic Scattering*, Nucl. Sci. Eng., 102,1 1989

[20] S. Oliveira, "Parallel Multigrid Methods for Transport Equations ", Ph. D. Thesis, Department of Mathematics, University of Colorado at Denver, Denver, Co., May, 1993.