An angular multigrid method for computing mono-energetic particle beams in Flatland

Christoph Börgers and Scott MacLachlan

Department of Mathematics, Tufts University, Medford, Massachusetts 02155

Abstract: Beams of microscopic particles penetrating scattering background matter play an important role in several applications. The parameter choices made here are motivated by the problem of electron-beam cancer therapy planning. Mathematically, a steady particle beam penetrating matter, or a configuration of several such beams, is modeled by a boundary value problem for a Boltzmann equation. Grid-based discretization of such a problem leads to a system of algebraic equations. This system is typically very large because of the large number of independent variables in the Boltzmann equation — six if no dimension-reducing assumptions other than time independence are made. If grid-based methods are to be practical for these problems, it is therefore necessary to develop very fast solvers for the discretized problems. For beams of mono-energetic particles interacting with a passive background, but not with each other, in two space dimensions, the first author proposed such a solver, based on angular domain decomposition, in [3]. Here, we propose and test an angular multigrid algorithm for the same model problem. Our numerical experiments show rapid, grid-independent convergence. For high-resolution calculations, our method is substantially more efficient than the angular domain decomposition method of [3]. In addition, unlike angular domain decomposition, the angular multigrid method works well even when the angular diffusion coefficient is fairly large.

1   Introduction

Charged-particle transport plays an important role in many fields; examples include electron microscopy [20], cancer therapy using electrons [12, 13], protons, or heavy ions [22], and various other applications of ion beams [21, 26]. The work presented here aims to contribute to the development of accurate and efficient simulation methods for charged-particle transport. The parameter choices in this paper are motivated by the electron-beam cancer therapy dose calculation problem [12, 13]. Procedures for electron-beam cancer treatment plan optimization require the solution of many electron transport problems; the efficiency of the algorithms used for these transport problems is therefore important.

Mathematically, a particle beam, or a configuration of several such beams, is modeled by a Boltzmann equation. This equation may be linear or nonlinear, depending on whether or not the beam particles interact with each other. Here, we will assume linearity, a common and accurate approximation in electron-beam cancer therapy planning.

Thus, our investigation belongs to the vast subject of numerical methods for the linear Boltzmann equation. One important source of difficulty in the computational solution of the linear Boltzmann equation is the sheer size of the problems: There are, in general, seven independent variables (position and velocity in three dimensions, and time), and still six when considering time-independent boundary value problems, as we do here. Of course,

this is the number of independent variables in any kinetic problem (unless the geometry is special), not just in charged-particle transport. However, there are additional difficulties associated specifically with charged-particle transport: The mean free path tends to be small, scattering tends to be very *forward-peaked* (*i.e.*, particles are typically deflected only very slightly by a single interaction with the background), and particles typically lose very little energy in a single interaction. These properties of charged-particle transport cause difficulties with the accuracy of discretizations and with the efficiency of solution algorithms for the discretized problems [16, Section 3.2], which have lead many in the Medical Physics community to believe that the most efficient way of modeling electron beams may be Monte Carlo simulation. However, based on a rough theoretical complexity estimate presented in [4], we believe that deterministic, grid-based methods could eventually prove to be a very attractive alternative to Monte Carlo simulation, provided that all available tools of numerical computing are brought to bear to develop highly accurate discretizations as well as optimally efficient solution algorithms for the discretized problems. Some algorithm and code development efforts in this direction are in fact underway; see, for instance, [2] and [9].

In this paper, we focus on the problem of designing highly efficient solvers for a grid-based discretization of the model equation of [3] (reviewed in Section 2). This equation describes physics in "Flatland" [1], *i.e.*, in a fictitious two-dimensional world. It is arguably the simplest possible caricature of charged-particle transport in more than one space dimensions. We propose and test an angular multigrid method for this problem. The idea of angular multigrid methods for particle transport with forward-peaked scattering was first proposed, for a one-dimensional problem, by J. Morel and T. Manteuffel [17]; extensions of the idea to higher dimensions have had limited success so far [18]. However, for our two-dimensional model problem, the convergence of the angular multigrid method turns out to be rapid, and the speed of convergence appears to be independent of the grid size. It is not entirely clear at this point why our approach, for the simpler problem discussed here, does not enounter the obstacles described in [18]; see Section 10 (Discussion) for some thoughts on this point.

2 The Model Problem

To make this paper as self-contained as possible, we will review the model equation of [3] and its properties here, closely following but abbreviating the exposition of [3]. We will mix physical and mathematical terminology, writing, for instance, about "particles" that move "in a domain $\Omega \subseteq \mathbb{R}^2$".

2.1 Model equation

We consider mono-energetic particle transport in two space dimensions. It must be emphasized that this is *not* the same as (and, indeed, is simpler than) the projection of three-dimensional particle transport into a plane. We consider the motion of particles in a domain, $\Omega \subseteq \mathbb{R}^2$, assuming that all particles move at the same constant speed, $c > 0$. Each particle experiences collisions at random times, causing random direction changes. The inter-

collision distances, $\lambda > 0$, are exponentially distributed and independent of each other; their expectation, $\bar{\lambda} > 0$, is called the *mean free path*. The deflection angles, η, (see Figure 1) are independent of each other and of the inter-collision distances.
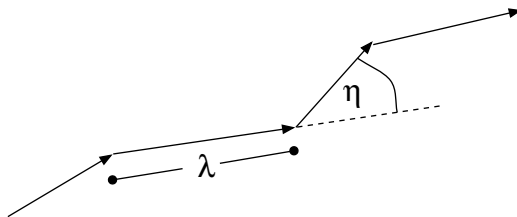

Figure 1: *An example of a particle path.*

The probability density of η is $p : (-\pi, \pi) \to \mathbb{R}_+$. We assume that p is an even function, *i.e.*, that particles have no preference for scattering to the right over scattering to the left or vice versa. The graph of p qualitatively looks like that shown in Figure 2, where the forward-peakedness of the scattering is reflected by the peak in the graph of p near $\eta = 0$. For realistic models of the scattering of electrons, however, this peak would be much more pronounced than in Figure 2.
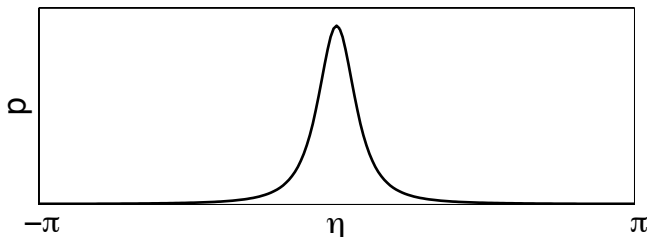

Figure 2: *A qualitative plot of the probability density p.*

The *phase space density*, $f = f(x, y, \theta, t)$, is the number of particles per unit (x, y, θ)-volume, where $(x, y) \in \Omega$ denotes the particle position, $(\cos\theta, \sin\theta)$ is the particle direction, and $t \geq 0$ is time. The function f is 2π-periodic in θ. The time evolution of f is governed by the *linear Boltzmann equation*, the mathematical statement of the law of conservation of particles,

$$f_t + c\cos\theta f_x + c\sin\theta f_y = cQf \ . \tag{1}$$

Here, c denotes the particle speed (assumed constant in this model problem), and the *collision operator Q* is defined by

$$Qf = \frac{p * f - f}{\bar{\lambda}} \ ,$$

where $*$ denotes convolution with respect to θ:

$$(p * f)(\theta) = \int_{-\pi}^{\pi} p(\eta) f(\theta - \eta) \, d\eta \ .$$

As mentioned earlier, in some applications, including electron-beam radiation therapy, the primary interest is in steady-state problems. We therefore drop the term f_t in Eq. (1). We then also divide both sides by c. The result is

$$\cos\theta \, f_x + \sin\theta \, f_y = \frac{p * f - f}{\bar{\lambda}} \ . \tag{2}$$

This is our model equation.

We define the Fourier coefficients,

$$\hat{p}_n = \int_{-\pi}^{\pi} \exp(-in\eta) p(\eta) \, d\eta \tag{3}$$

and

$$\hat{f}_n = \frac{1}{2\pi} \int_{-\pi}^{\pi} \exp(-in\tau) f(\tau) \, d\tau \tag{4}$$

for all integers n (where, for now, we write $f = f(\theta)$ and suppress the dependence on (x, y)). Note that we include the factor of $1/(2\pi)$ in (4) but not in (3); this is intentional and will simplify the notation a bit. With these definitions, we have

$$f(\theta) = \sum_{n=-\infty}^{\infty} \hat{f}_n \exp(in\theta) \ ,$$

and it is straightforward to verify that

$$Qf(\theta) = \sum_{n=-\infty}^{\infty} \frac{\hat{p}_n - 1}{\bar{\lambda}} \hat{f}_n \exp(in\theta). \tag{5}$$

This equation serves as the basis of our discretization of Q; see Section 3. In analyzing this discretization, the following observations about $\hat{p}_n$ will be useful. Since p is assumed to be an even function,

$$\hat{p}_n = \int_{-\pi}^{\pi} \cos(n\eta) \, p(\eta) \, d\eta \ .$$

This implies that

$$\hat{p}_0 = 1, \quad -1 < \hat{p}_n < 1 \quad \text{for all } n \neq 0, \text{ and} \quad \hat{p}_{-n} = \hat{p}_n \quad \text{for all } n. \tag{6}$$

When collisions are strongly forward-peaked (that is, when $p(\eta)$ is small everywhere except near $\eta = 0$), the following calculation is plausible:

$$\frac{p * f - f}{\bar{\lambda}}(\theta) = \frac{1}{\bar{\lambda}} \left(\int_{-\pi}^{\pi} p(\eta) \, f(\theta - \eta) \, d\eta - f(\theta) \right) \approx$$

$$\frac{1}{\bar{\lambda}} \left(\int_{-\pi}^{\pi} p(\eta) \left(f(\theta) - f_\theta(\theta)\eta + f_{\theta\theta}(\theta) \frac{\eta^2}{2} \right) d\eta - f(\theta) \right) = D f_{\theta\theta}(\theta) \ ,$$

where the *angular diffusion coefficient*, D, is defined by

$$D = \frac{1}{2\bar{\lambda}} \int_{-\pi}^{\pi} \eta^2 p(\eta) \, d\eta \ . \tag{7}$$

4

This calculation was made rigorous in [3] as follows. Assume that

$$\bar{\lambda} \to 0 \quad \text{and} \quad \int_{-\pi}^{\pi} \eta^2 p(\eta)\, d\eta \to 0 \,,$$

that is, collisions become increasingly frequent and increasingly forward-peaked. Assume further that there is a balance between these two limits in the following sense:

$$\frac{\int_{-\pi}^{\pi} \eta^2 p(\eta)\, d\eta}{2\bar{\lambda}} \to D > 0 \,.$$

We then ask whether

$$Qf \rightharpoonup D \frac{\partial^2 f}{\partial^2 \theta} \,, \tag{8}$$

where "$\rightharpoonup$" denotes weak convergence in L^2 (that is, convergence of Fourier coefficients). Proposition 1 of [3] states that (8) holds if and only if

$$\frac{\int_{-\pi}^{\pi} \eta^4 p(\eta)\, d\eta}{\int_{-\pi}^{\pi} \eta^2 p(\eta)\, d\eta} \to 0 \,. \tag{9}$$

We now turn to the question of how to construct a probability density, p, that mimics realistic three-dimensional scattering kernels as far as possible in Flatland. It is natural to consider $p \sim 1/|\eta|^q$ for some $q > 0$. To obtain a 2π-periodic function that is non-singular at $\eta = 0$, we replace $|\eta|$ by $\sqrt{2(1 - \cos\eta) + \varepsilon^2}$ for some small parameter $\varepsilon > 0$; see Figure 3. Thus we define

$$p(\eta) = \frac{C}{(2(1 - \cos\eta) + \varepsilon^2)^{q/2}} \,, \tag{10}$$

with the constant $C > 0$ chosen so that $\int_{-\pi}^{\pi} p(\eta)\, d\eta = 1$.
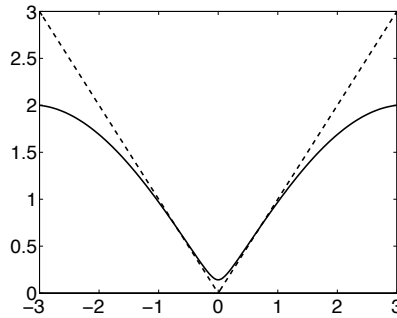

Figure 3: $|\eta|$ *(dashes) vs.* $\sqrt{2(1 - \cos\eta) + \varepsilon^2}$ *(solid curve) with* $\varepsilon = 0.1$.

Proposition 2 of [3] shows that, for this choice of $p(\eta)$,

$$\lim_{\varepsilon \to 0} \int_{-\pi}^{\pi} \eta^2 p(\eta)\, d\eta = 0 \quad \text{if and only if } q \geq 1 \,,$$

and

$$\lim_{\varepsilon \to 0} \frac{\int_{-\pi}^{\pi} \eta^4 p(\eta)\, d\eta}{\int_{-\pi}^{\pi} \eta^2 p(\eta)\, d\eta} = 0 \quad \text{if and only if } q \geq 3 \,.$$

5

Thus, for $q \geq 3$, Eq. (2) can be approximated by the Fokker-Planck equation,

$$\cos\theta f_x + \sin\theta f_y = D f_{\theta\theta}, \tag{11}$$

when $\bar{\lambda}$ and $\int_{-\pi}^{\pi} \eta^2 p(\eta)\, d\eta$ are small, with D defined as in (7).

In the borderline case, $q = 3$, the convergence in (9) is logarithmic [3]. In three space dimensions, the screened Rutherford scattering cross section [25] has the precisely analogous properties [5, Appendix]. We therefore say that (10) with $q = 3$ defines *screened Rutherford scattering in Flatland*.

Another interesting case is $q = 2$. In this case, the approximation of Q by $D\partial^2/\partial\theta^2$ is not valid. In fact, it is not hard to show that in the limit as $\varepsilon \to 0$, one obtains a pseudo-differential operator proportional to $-\sqrt{-\partial^2/\partial\theta^2}$. In this and some other regards, (10) with $q = 2$ resembles the Henyey-Greenstein scattering cross section [11]; see [19] and [3]. We therefore say that (10) with $q = 2$ defines *Henyey-Greenstein scattering in Flatland*.

In the examples of this paper, we always specify the parameters q, $\bar{\lambda}$, and D, not the less intuitive parameter ε. Note that ε can be calculated, using Eqs. (7) and (10), from q, $\bar{\lambda}$ and D.

2.2 Boundary conditions

We supplement Eq. (2) or its Fokker-Planck approximation (11) with *inflow boundary conditions*. That is, $f(x, y, \theta)$ is prescribed if (x, y) lies on the boundary, $\partial\Omega$, of Ω and the direction vector, $(\cos\theta, \sin\theta)$, points from (x, y) into Ω. The details of how this is implemented in a discrete approximation of the continuous boundary value problem turn out to be crucial for the efficiency of our multigrid algorithm; see Section 4.1.

2.3 Choice of parameters

The choice of the parameters D and $\bar{\lambda}$ was discussed in detail in [3, Section 5]. It was shown there that, approximately, the amount by which a beam broadens as it penetrates a strip of width L in the (x, y)-plane equals $(DL)^{1/2} L$. For the beam to remain recognizable as a beam, $(DL)^{1/2} L$ should be much smaller than L, *i.e.*, $\sqrt{DL} \ll 1$. It was further argued in [3, Section 5] that the values of $L/\bar{\lambda}$ that are most relevant in electron-beam radiation therapy are on the order of (a few multiples of) $10,000$.

For simplicity, we consider $\Omega = (0, L) \times (0, L)$ throughout this paper, and assume that x and y are non-dimensionalized so that L becomes 1:

$$\Omega = (0, 1) \times (0, 1) . \tag{12}$$

The condition $\sqrt{DL} \ll 1$ then becomes

$$D \ll 1 .$$

The condition $L/\bar{\lambda} \sim 10,000$ becomes $\bar{\lambda} \sim 1/10,000$. Many of our numerical results are for $D = 0.1$ and $\bar{\lambda} = 1/20,000$, but in Section 9, we will also test much larger values of D, and a broad range of values of $\bar{\lambda}$.

2.4 The spatial diffusion limit

Very large values of D may be of lesser interest in electron-beam radiation therapy. However, we will examine the behavior of our multigrid algorithm for larger values of D in Section 9.1.4. To elucidate the nature of the problem for large D, we will briefly consider the limit $D \to \infty$ here, for simplicity focusing exclusively on the Fokker-Planck equation (11). In this limit, our model transport process turns into diffusion in the (x, y)-plane. This is entirely analogous to the standard asymptotic expansion of the neutron transport equation in the limit of vanishing mean free path [10, 14]; in fact, $D \to \infty$ implies $\bar{\lambda} \to 0$ by Eq. (7). We present here the formal derivation of the spatial (steady-state) diffusion equation from Eq. (11) in the limit $D \to \infty$.

Suppose that

$$D = \frac{D_0}{\delta} ,$$

with $D_0 > 0$ fixed, $\delta > 0$, $\delta \to 0$. Assume an asymptotic expansion of the form

$$f = f^{(0)} + \delta f^{(1)} + \delta^2 f^{(2)} + ..., \tag{13}$$

where $f^{(0)}$, $f^{(1)}$, $f^{(2)}$, ... are functions of (x, y, θ), but not of δ, and are periodic in θ with period 2π. We insert (13) into (11):

$$\cos\theta \left(f^{(0)} + \delta f^{(1)} + \delta^2 f^{(2)} + ... \right)_x + \sin\theta \left(f^{(0)} + \delta f^{(1)} + \delta^2 f^{(2)} + ... \right)_y =$$

$$\frac{D_0}{\delta} \left(f^{(0)} + \delta f^{(1)} + \delta^2 f^{(2)} + ... \right)_{\theta\theta} . \tag{14}$$

Isolating terms containing the factor δ^{-1}, we find

$$f^{(0)}_{\theta\theta} = 0 .$$

Thus $f^{(0)}$ is a linear function of θ. Since $f^{(0)}$ is also 2π-periodic in θ, we conclude

$$f^{(0)} = f^{(0)}(x, y) . \tag{15}$$

Isolating terms containing the factor δ^0 in (14), we now find

$$\cos\theta f^{(0)}_x + \sin\theta f^{(0)}_y = D_0 f^{(1)}_{\theta\theta} . \tag{16}$$

Integrating twice, we obtain

$$D_0 f^{(1)} = -\cos\theta f^{(0)}_x - \sin\theta f^{(0)}_y + a(x, y) + b(x, y)\theta .$$

Since $f^{(1)}$ is periodic in θ, we must have $b(x, y) \equiv 0$:

$$D_0 f^{(1)} = -\cos\theta f^{(0)}_x - \sin\theta f^{(0)}_y + a(x, y) . \tag{17}$$

Isolating terms containing the factor δ^1 in Eq. (14), we obtain

$$\cos\theta f^{(1)}_x + \sin\theta f^{(1)}_y = D_0 f^{(2)}_{\theta\theta} . \tag{18}$$

7

Inserting Eq. (17) into (18) and integrating over θ, we find Laplace's equation for $f^{(0)}$:

$$f^{(0)}_{xx} + f^{(0)}_{yy} = 0 . \tag{19}$$

Thus for large D, a solution of (11) will be nearly isotropic (independent of θ) in the interior of the domain Ω, and satisfy Laplace's equation. The solution will not be isotropic on the boundary, however; there will be a boundary layer in which a rapid transition from anisotropy on the boundary to isotropy in the interior will occur.

3   Discretization in angle

Following [3], we discretize the collision operator, Q, based on Eq. (5):

$$Qf(\theta) = \frac{p * f - f}{\bar{\lambda}} = \sum_{n=-\infty}^{\infty} \frac{\hat{p}_n - 1}{\bar{\lambda}} \hat{f}_n \exp(in\theta) =$$

$$\sum_{n=-\infty}^{\infty} \frac{\hat{p}_n - 1}{\bar{\lambda}} \frac{1}{2\pi} \int_{-\pi}^{\pi} \exp(-in\tau) f(\tau) \, d\tau \, \exp(in\theta) . \tag{20}$$

We choose a positive integer, n_θ, divisible by 2 for simplicity, and define

$$\Delta\theta = \frac{2\pi}{n_\theta} .$$

The discretization of Q is an operator, $Q^{\Delta\theta}$, that maps the space of 2π-periodic functions defined on the grid

$$\Gamma_{n_\theta} = \{\theta_l = l\Delta\theta \ : \ l \text{ integer}\}$$

into itself. A 2π-periodic function, f, on Γ_{n_θ} can naturally be identified with the vector

$$\mathbf{f} = (f(\theta_l))_{l=-n_\theta/2+1, -n_\theta/2+1, \dots, n_\theta/2} .$$

Therefore $Q^{\Delta\theta}$ can also be thought of as an $n_\theta \times n_\theta$-matrix. Using this point of view, $Q^{\Delta\theta}$ is defined by the following discrete analog of Eq. (20):

$$\left(Q^{\Delta\theta}\mathbf{f}\right)_l = \sum_{n=-n_\theta/2+1}^{n_\theta/2} \frac{\hat{p}_n - 1}{\bar{\lambda}} \frac{1}{n_\theta} \sum_{m=-n_\theta/2+1}^{n_\theta/2} \exp(-in\theta_m) f(\theta_m) \exp(in\theta_l) . \tag{21}$$

An equivalent formula is obtained by replacing n with $-n$ in the summands on the right-hand side of Eq. (21):

$$\left(Q^{\Delta\theta}\mathbf{f}\right)_l = \sum_{n=-n_\theta/2+1}^{n_\theta/2} \frac{\hat{p}_{-n} - 1}{\bar{\lambda}} \frac{1}{n_\theta} \sum_{m=-n_\theta/2+1}^{n_\theta/2} \exp(in\theta_m) f(\theta_m) \exp(-in\theta_l) . \tag{22}$$

Averaging Eqs. (21) and (22) and using $\hat{p}_{-n} = \hat{p}_n$ (see (6)), we find

$$\left(Q^{\Delta\theta}\mathbf{f}\right)_l = \frac{1}{n_\theta} \sum_{n=-n_\theta/2+1}^{n_\theta/2} \sum_{m=-n_\theta/2+1}^{n_\theta/2} \frac{\hat{p}_n - 1}{\bar{\lambda}} \cos(n(\theta_l - \theta_m)) f(\theta_m) . \tag{23}$$

8

This form of the equation shows that $Q^{\Delta\theta}$ is real and symmetric. The eigenvalues of $Q^{\Delta\theta}$ are $(\hat{p}_n - 1)/\overline{\lambda}$, $-n_\theta/2 + 1 \le n \le n_\theta/2$, with associated eigenvectors

$$(\exp(in\theta_l))_{-n_\theta/2+1 \le l \le n_\theta/2}$$

[3, Section 10]. Since $\hat{p}_0 = 1$ and $|\hat{p}_n| < 1$ for all $n \ne 0$ (see (6)), we conclude that $Q^{\Delta\theta}$ is negative semi-definite with a one-dimensional kernel spanned by vector $[1, 1, ..., 1]^T \in \mathbb{R}^{n_\theta}$; this is Proposition 4 of [3].

In the Fokker-Planck limit,

$$\frac{\hat{p}_n - 1}{\overline{\lambda}} \to -Dn^2 \; ;$$

this is Eq. (14) of [3]. In this limit, the right-hand side of Eq. (21) therefore becomes

$$\sum_{n=-n_\theta/2+1}^{n_\theta/2} (-Dn^2) \left(\frac{1}{n_\theta} \sum_{m=-n_\theta/2+1}^{n_\theta/2} \exp(-in\theta_m) f(\theta_m) \right) \exp(in\theta_l) \; . \tag{24}$$

Note that $-n^2$ is the eigenvalue of $\partial^2/\partial\theta^2$ associated with the eigenfunction $\exp(in\theta)$. Eq. (24) represents a spectrally accurate discretization of $Df_{\theta\theta}$. We will also report on numerical experiments using the standard second-order three-point discretization of the second derivative with respect to θ on the right-hand side of the Fokker-Planck equation (11).

4 Discretization in space

In this paper, it is not our aim to address the question how to best discretize our model equation in space. In what follows, we propose a relaxation technique and an associated multigrid algorithm that we believe will be efficient for many reasonable discretizations. The central assumptions of this method are that the discretization of the scattering term (either in the full scattering operator or its Fokker-Planck approximation) preserves the elliptic character of this term, and that the spatial terms are, in some sense, discretized upstream, so that the effects of advection may be effectively computed using a downstream Gauss-Seidel-like iteration. Our focus in this paper is on the optimally efficient solution of the resulting discretized equations. We use what are arguably the simplest possible discretizations for the spatial terms, first- and second-order upstream differencing based on a uniform spatial grid.

4.1 The discrete inflow and outflow boundaries

Our spatial grid points are

$$(j\Delta s, k\Delta s), \quad 0 \le j, k \le n_s \; ,$$

where $n_s > 0$ is an integer and $\Delta s = 1/n_s$. As discussed in Section 3, the discrete values of θ are

$$\theta_l = l\Delta\theta, \quad -n_\theta/2 + 1 \le l \le n_\theta/2 \; ,$$

where $n_\theta > 0$ is an even integer and $\Delta\theta = 2\pi/n_\theta$.

It is important for the performance of the multigrid algorithm of Section 8 to be careful about what we mean, in the discrete case, by *inflow directions* for spatial grid points on the boundary of Ω. Let $(j\Delta s, k\Delta s)$ be a spatial boundary point, *i.e.*, $j = 0$, $j = n_s$, $k = 0$, or $k = n_s$. Let $-n_\theta/2 + 1 \leq l \leq n_\theta/2$. At first sight, it would seem natural to call $(\cos\theta_l, \sin\theta_l)$ an *inflow direction* if it points strictly into Ω from the boundary point $(j\Delta s, k\Delta s)$, and an *outflow direction* otherwise. This definition, however, is not quite compatible with standard definitions of grid-transfer operators, in a sense explained in Section 8.1. We have found that, as a result, multigrid convergence can deteriorate significantly as the angular grid is refined.

There are two different approaches to overcoming this problem. One is to modify standard grid-transfer and coarse-grid operators. Proper multigrid convergence can, in fact, be restored this way. However, we have adopted the simpler and more straightforward alternative of modifying the definitions of inflow and outflow directions, as described in the following paragraph. The modified definitions are compatible with standard grid-transfer operators, as explained in Section 8.1, and lead to grid-independent convergence of standard multigrid cycles, as demonstrated in Section 9.

At a point $(j\Delta s, k\Delta s) \in \partial\Omega$, a direction vector $(\cos\theta_l, \sin\theta_l)$ is called an *outflow direction* if the *negative* of $(\cos\theta_l, \sin\theta_l)$ points *strictly into* the domain. We also say then that $(j\Delta s, k\Delta s, l\Delta\theta)$ belongs to the *discrete outflow boundary*, denoted by $\partial\Omega^{out}$. (For simplicity, we do not indicate the dependence of $\partial\Omega^{out}$ on Δs and $\Delta\theta$ in this notation.) We call $(\cos\theta_l, \sin\theta_l)$ an *inflow direction* at the spatial boundary point $(j\Delta s, k\Delta s)$ if it is not an outflow direction, and we say then that $(j\Delta s, k\Delta s, l\Delta\theta)$ belongs to the *discrete inflow boundary*, denoted by $\partial\Omega^{in}$. Figure 4 illustrates the definitions of $\partial\Omega^{out}$ and $\partial\Omega^{in}$. (Note that in the corner points of $\partial\Omega$, many directions that would intuitively be considered outflow directions are in fact inflow directions by our definition.) The set

$$\Omega^\circ = \{(j\Delta s, k\Delta s, l\Delta\theta) \,|\, (j\Delta s, k\Delta s) \notin \partial\Omega\}$$

is called the *interior* of the grid. Thus, the spatial grid is partitioned into its interior and the inflow and outflow boundaries:

$$\{(j\Delta s, k\Delta s, l\Delta\theta) \,|\, 0 \leq j, k \leq n_s, \; -n_\theta/2 + 1 \leq l \leq n_\theta/2\} = \Omega^\circ \cup \partial\Omega^{out} \cup \partial\Omega^{in} \,.$$

Approximations

$$f_{j,k,l} \approx f(j\Delta s, k\Delta s, l\Delta\theta)$$

are prescribed if $(j\Delta s, k\Delta s, l\Delta\theta) \in \partial\Omega^{in}$, and computed from the finite-difference equations if $(j\Delta s, k\Delta s, l\Delta\theta) \in \partial\Omega^\circ \cup \partial\Omega^{out}$.

As indicated in Figure 4, we deviate from the naive definitions of inflow and outflow directions only in directions tangential to $\partial\Omega$, and in a range of directions in the corner points of $\partial\Omega$; thus our modifications affect only a set of measure zero in the continuous inflow boundary, and should not prevent convergence to the solution of the continuous problem as $\Delta s \to 0$ and $\Delta\theta \to 0$.

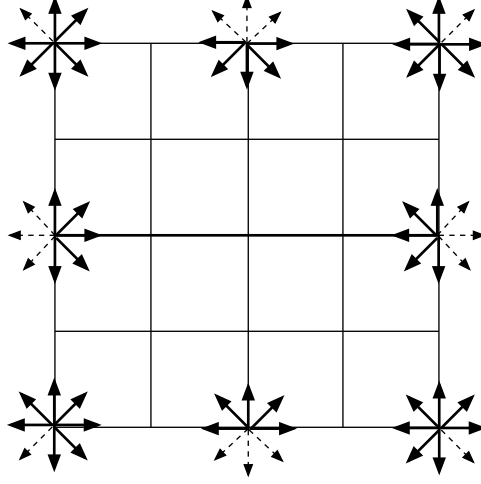Figure 4: *Discrete "inflow" (bold) and "outflow" (dashes) directions.*

4.2   First-order upstream differencing

The first-order upstream differencing method is based on the one-sided difference formula,

$$\varphi'(s) = \frac{\varphi(s) - \varphi(s - \Delta s)}{\Delta s} + O(\Delta s),$$

for the derivative, $\varphi'(s)$, of a smooth function, φ, of the real variable s. The difference quotients are taken in the upstream direction in all cases, that is, in the direction opposite to the direction vector $(\cos\theta, \sin\theta)$. As an example, suppose that $\cos\theta_l > 0$ and $\sin\theta_l < 0$; the first-order upstream discretization of

$$\cos\theta\, f_x + \sin\theta\, f_y \tag{25}$$

at $(x, y, \theta) = (j\Delta s, k\Delta s, l\Delta\theta)$ is then

$$\cos\theta_l \frac{f_{j,k,l} - f_{j-1,k,l}}{\Delta s} + \sin\theta_l \frac{f_{j,k+1,l} - f_{j,k,l}}{\Delta s}\ .$$

Note that this is well-defined, *i.e.*, $j - 1 \geq 0$ and $k + 1 \leq n_s$, if $(j\Delta s, k\Delta s, l\Delta\theta) \notin \partial\Omega^{in}$. The first-order upstream differencing approximation for (25) is, in general, well-defined for $(j\Delta s, k\Delta s, l\Delta\theta) \notin \partial\Omega^{in}$. At points $(j\Delta s, k\Delta s, l\Delta\theta) \in \partial\Omega^{in}$, the transport equation is not discretized, but instead $f_{j,k,l}$ is prescribed.

4.3   Second-order upstream differencing.

The second-order upstream differencing method is based on the one-sided difference formula

$$\varphi'(s) = \frac{1.5\varphi(s) - 2\varphi(s - \Delta s) + 0.5\varphi(s - 2\Delta s)}{\Delta s} + O(\Delta s^2)\ .$$

Again the difference quotients are taken in the upstream directions. As in the example above, suppose again that $\cos\theta_l > 0$ and $\sin\theta_l < 0$. The second-order upstream discretization of (25) is then

$$\cos\theta_l \frac{1.5f_{j,k} - 2f_{j-1,k} + 0.5f_{j-2,k}}{\Delta s} + \sin\theta_l \frac{-0.5f_{j,k+2} + 2f_{j,k+1} - 1.5f_{j,k}}{\Delta s} \ . \tag{26}$$

Note that this is not well-defined in points with $j = 1$ or $k = n_s - 1$. When the second-order discretization of f_x or f_y is not well-defined, we use the first-order discretization instead. For instance, when $j = 1$ but $k < n_s - 1$, then (26) is replaced by

$$\cos\theta_l \frac{f_{j,k} - f_{j-1,k}}{\Delta s} + \sin\theta_l \frac{-0.5f_{j,k+2} + 2f_{j,k+1} - 1.5f_{j,k}}{\Delta s} \ .$$

Since first-order discretizations are used only in grid points immediately adjacent to the boundary, we expect the overall order of convergence of our discretization scheme to be two; the numerical experiments of Table 1 in Section 6 support this.

5 Test problems

In what follows, we will consider three test problems. The first two are chosen to model somewhat realistic broad beams of electrons in two distinct configurations. The third test problem is chosen specifically to test the numerical performance of the proposed algorithms.

5.1 Test problem 1: A broad beam aligned with the grid

We first consider a broad beam in which the mean direction of incoming particles is aligned with the grid. The inflow boundary condition is of the form

$$f(x,y,\theta) = \begin{cases} g(x)h(\theta \bmod 2\pi) & \text{if } 0 \le x \le 1, \ y = 0, \ 0 \le \theta \bmod 2\pi \le \pi , \\ 0 & \text{otherwise}, \end{cases} \tag{27}$$

where

$$\theta \bmod 2\pi = \theta + 2\pi u, \quad u = \text{integer chosen so that } -\pi < \theta \bmod 2\pi \le \pi .$$

The function $g = g(x)$ is a smoothed step function; its graph is depicted in Figure 5. The formulas defining g are

$$g(x) = \begin{cases} 0 & \text{if } 0 \le x \le 0.2 , \\ (1 + \tanh(2\tan(5(x-0.3)\pi)))/2 & \text{if } 0.2 \le x \le 0.4 , \\ 1 & \text{if } 0.4 \le x \le 0.6 , \\ (1 - \tanh(2\tan(5(x-0.7)\pi)))/2 & \text{if } 0.6 \le x \le 0.8 , \\ 0 & \text{if } 0.8 \le x \le 1 . \end{cases} \tag{28}$$

(This is an infinitely often differentiable function.) We further define

$$h(\theta) = \frac{e^{-(\theta-\pi/2)^2/0.1}}{\sqrt{0.1\pi}} \ . \tag{29}$$
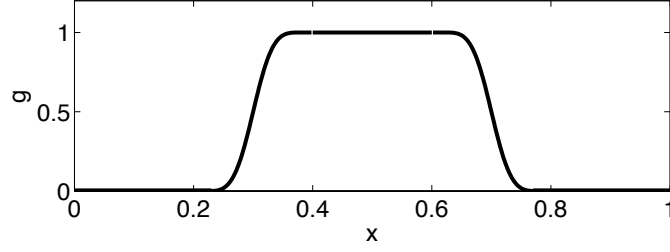
Figure 5: *Spatial profile of the beams in test problems 1 and 2 at the inflow boundary.*
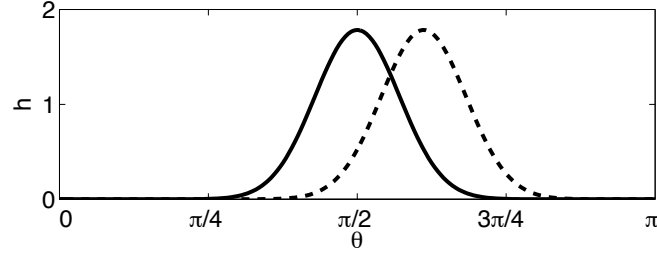

Figure 6: *Angular profile of the beams in test problems 1 and 2 (solid and dashed) at the inflow boundary.*

Thus the mean angle at which particles enter is $\pi/2$. The solid curve in Figure 6 is the graph of (29). Figure 7A shows the *macroscopic density*,

$$\rho_{j,k} = \sum_{l=-n_\theta/2+1}^{n_\theta/2} f_{j,k,l} \, \Delta\theta,$$

for this example, using screened Rutherford scattering with $D = 0.1$ and $\bar{\lambda} = 1/20,000$, displayed on a 32×32-grid, but computed using second-order upstream differencing on a $512 \times 512 \times 512$-grid.
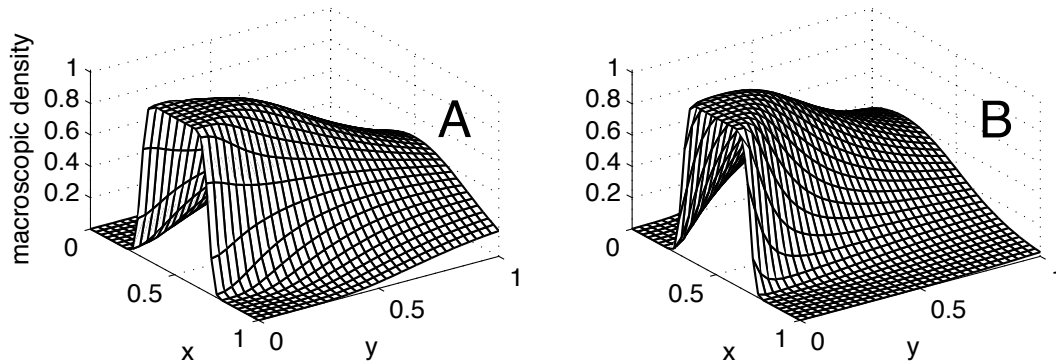

Figure 7: *Macroscopic density for two broad beams, one (A) aligned with the grid, and the other (B) not.*

5.2 Test problem 2: A broad beam not aligned with the grid

Our second test problem differs from the first only in the definition of the function h. In comparison with the first test problem, we shift the mean incoming particle direction by 0.35 radians, or approximately $20°$, so that

$$h(\theta) = \frac{e^{-(\theta - \pi/2 - 0.35)^2/0.1}}{\sqrt{0.1\pi}} \ . \tag{30}$$

The dashed line in Figure 6 is the graph of h. The resulting macroscopic density is shown in Figure 7B. We expect (and will demonstrate numerically) that our spatial discretizations are somewhat less accurate for test problem 2 than for test problem 1.

5.3 Test problem 3: Zero inflow boundary values

To measure asymptotic convergence factors of the proposed iterative methods, it is useful to be able to carry out many multigrid iterations without encountering round-off effects. We, therefore, will also use a test problem in which — as in test problems 1 and 2 — there are no interior sources, but in which there is also no inflow through the boundary, so that the solution is zero:

$$f_{j,k,l} = 0 \quad \text{for all } j, k, \text{ and } l \ .$$

We then start our iterations (see Sections 7–9) with a nonzero initial guess (see Eq. (37)), in order to more accurately measure asymptotic performance of the resulting iterations.

6 Accuracy of the discretizations

The accuracy of the finite-difference discretizations is not the principal topic of this paper; we are mainly concerned with the efficient solution of the discretized problems here. Nevertheless, in this subsection, we report on some numerical experiments concerning the accuracy of our discretizations. The main purpose of these experiments is to verify that first- and second-order convergence are in fact seen with the first- and second-order upstream differencing schemes.

The experiments of this subsection are for test problems 1 and 2 with screened Rutherford scattering, using $D = 0.1$ and $\overline{\lambda} = 1/20,000$, and using first and second order upstream differencing. Let

$$f_{j,k,l}^{(n)} , \quad 0 \le j, k \le n, \quad -n/2 + 1 \le l \le n/2,$$

denote the approximations computed on the grid with $n_s = n_\theta = n$. For $n = 2^r \times 32$, $r \ge 0$ integer, we define $E^{(n)}$ to be the mean square discrepancy between $f^{(n)}$ and $f^{(2n)}$, both restricted to the $32 \times 32 \times 32$-grid:

$$E^{(n)} = \frac{1}{33^2 \cdot 32} \sum_{j=0}^{32} \sum_{k=0}^{32} \sum_{l=-15}^{16} \left(f_{2^r j, 2^r k, 2^r l}^{(n)} - f_{2^{r+1} j, 2^{r+1} k, 2^{r+1} l}^{(2n)} \right)^2 \ . \tag{31}$$

We compute $f^{(n)}$ for $n = 32, 64, 128, 256,$ and 512, $E^{(n)}$ for $n = 32, 64, 128,$ and 256, and $R^{(n)} = E^{(n)}/E^{(2n)}$ for $n = 32, 64,$ and 128. Table 1 shows the results. As expected, discretization accuracy is better for test problem 1 than for test problem 2, but the difference is not dramatic. As n increases, the ratio $R^{(n)}$ appears to approach 2.0 for the first-order method, and 4.0 for the second-order method for both test problems, as it should.

	first order				second order			
	test problem 1		test problem 2		test problem 1		test problem 2	
	$E^{(n)}$	$R^{(n)}$	$E^{(n)}$	$R^{(n)}$	$E^{(n)}$	$R^{(n)}$	$E^{(n)}$	$R^{(n)}$
$n = 32$	0.0062	1.7	0.010	1.6	0.0039	2.7	0.0056	2.6
$n = 64$	0.0037	1.7	0.0065	1.7	0.0014	3.4	0.0021	3.4
$n = 128$	0.0021	1.8	0.0040	1.8	0.00042	3.8	0.00062	3.9
$n = 256$	0.0011	—	0.0023	—	0.00011	—	0.00016	—

Table 1: $E^{(n)} =$ *approximate mean square error (see (31)), and* $R^{(n)} = E^{(n)}/E^{(2n)}$

7   An angular relaxation scheme

The discretization of (2) that we wish to solve can briefly be represented as follows:

$$\left(\cos \theta_l \, \partial_x^{\Delta s} + \sin \theta_l \, \partial_y^{\Delta s} \right) f = Q^{\Delta \theta} f \, , \tag{32}$$

where $f = (f_{j,k,l})_{0 \leq j,k \leq n_s, -n_\theta/2+1 \leq l \leq n_\theta/2}$, the symbols $\partial_x^{\Delta s}$ and $\partial_y^{\Delta s}$ denote upstream discretizations of the partial derivatives with respect to x and y (see Sections 4.2 and 4.3), and $Q^{\Delta \theta}$ denotes the spectral discretization of the collision operator Q (see Section 3) or the three-point discretization of the Fokker-Planck operator $D\partial^2/\partial\theta^2$. Eq. (32) is supplemented with inflow boundary conditions (see Section 4.1).

In Section 8, we propose a multigrid correction cycle [6] to solve this problem, using coarsening in θ only, not in x and y. The auxiliary equations solved on coarser grids are of the form

$$\left(\cos \Theta_l \partial_x^{\Delta s} + \sin \Theta_l \partial_y^{\Delta s} \right) f = Q^{2^k \Delta \theta} f + R \, , \tag{33}$$

where $k \geq 1$ is an integer, and R denotes residuals transferred from the next finer grid. Eq. (33) is supplemented with zero inflow boundary conditions. We write

$$\Delta \Theta = 2^k \Delta \theta$$

to denote the mesh width of the angular grid, and

$$n_\Theta = \frac{n_\theta}{2^k}$$

to denote the number of angles that belong to the grid. We will always assume that n_Θ is even. The grid points are

$$\Theta_l = l \Delta \Theta, \quad -n_\Theta/2 + 1 \leq l \leq n_\Theta/2 \, .$$

15

We now describe a relaxation scheme for (33), supplemented with discrete inflow boundary conditions. In the multigrid cycle, this relaxation scheme is used on the finest grid (where R is zero, and the inflow boundary values are typically nonzero) as well as on the coarser grids (where R is typically nonzero, and the inflow boundary values are zero). We use the following "red-black" ordering of the angular grid points in the relaxation scheme:

$$l = -n_\Theta/2+1, -n_\Theta/2+3, ..., n_\Theta/2-1, \quad -n_\Theta/2+2, -n_\Theta/2+4, ..., n_\Theta/2 . \quad (34)$$

In several contexts, red-black (or black-red) ordering has been found to lead to significantly more rapid multigrid convergence than lexicographic ordering [23].

For each fixed l, we sweep over the spatial grid points in the downstream direction. At the end of such a sweep, for the fixed value of l, all residuals at grid points $(j\Delta s, k\Delta s, l\Delta\Theta)$ are zero. Of course, they typically become nonzero again later, as a result of sweeping over grid points with different values of l. The relaxation scheme is analogous to the waveform relaxation method for the heat equation discussed by Vandewalle and Horton [24]. Here, instead of a single "time-like" variable, we consider planes of (x,y) values for each Θ_l and perform a red-black relaxation sweep over these planes, combined with coarse-grid correction over the Θ-direction in our computational mesh.

7.1 Measuring convergence speed

In the following sections, we report on tests of the speed with which the proposed iterative methods converge. In each iteration, we determine the factor by which the sum of the squares of residuals is reduced. We denote this factor by ρ_ν for the ν-th iteration. We also use the notation

$$\bar{\rho}_{M,N} = \left(\prod_{\nu=M}^{N} \rho_\nu \right)^{1/(N-M+1)} \quad (35)$$

for the average convergence factor in cycles M through N; here $N \geq M \geq 1$.

The value of $\bar{\rho}_{M,N}$ depends, of course, on the initial guess. For test problems 1 and 2, we always use the initial guess

$$f_{j,k,l} = \begin{cases} \text{given boundary values} & \text{if } (j\Delta s, k\Delta s, l\Delta\theta) \text{ is an inflow boundary point,} \\ 0 & \text{otherwise.} \end{cases} \quad (36)$$

For test problem 3, we use the initial guess

$$f_{j,k,l} = \begin{cases} 1 & \text{if } 1 \leq j,k \leq n-1 , \\ 0 & \text{otherwise .} \end{cases} \quad (37)$$

7.2 Performance of relaxation

Tables 2 and 3 show the averaged per-cycle convergence factors, $\bar{\rho}_{1,24}$, for the Flatland analog of screened Rutherford scattering ($q = 3$ in Eq. (10)) with $D = 0.1$ and $\bar{\lambda} = 1/20,000$, using first- and second-order upstream differencing, respectively. For fixed n_s, the value

16

	$n_\theta = 32$	$n_\theta = 64$	$n_\theta = 128$	$n_\theta = 256$	$n_\theta = 512$
$n_s = 32$	0.315	0.667	0.938	0.965	0.982
$n_s = 64$	0.307	0.657	0.932	0.959	0.974
$n_s = 128$	0.302	0.647	0.925	0.952	0.967
$n_s = 256$	0.297	0.638	0.915	0.944	0.959
$n_s = 512$	0.292	0.629	0.905	0.935	0.951

Table 2: *Average convergence factor per iteration in first 24 relaxation iterations for test problem 2, discretized using first-order upstream differencing.*

	$n_\theta = 32$	$n_\theta = 64$	$n_\theta = 128$	$n_\theta = 256$	$n_\theta = 512$
$n_s = 32$	0.304	0.662	0.953	0.977	0.988
$n_s = 64$	0.299	0.652	0.943	0.971	0.984
$n_s = 128$	0.294	0.642	0.930	0.963	0.977
$n_s = 256$	0.290	0.633	0.917	0.951	0.969
$n_s = 512$	0.286	0.624	0.903	0.938	0.958

Table 3: *Average convergence factor per iteration in first 24 relaxation iterations for test problem 2, discretized using second-order upstream differencing.*

of $\bar{\rho}_{1,24}$ increases significantly with n_θ. For fixed n_θ, it appears that $\bar{\rho}_{1,24}$ can be bounded independently of n_s.

The results of Tables 2 and 3 are not surprising. The relaxation algorithm chosen here is based on successive sweeps across the *xy*-planes of the grid. Since such a sweep (temporarily) removes residuals in the plane regardless of n_s, it is to be expected that the convergence factors show little dependence on n_s. The slight decrease in $\bar{\rho}_{1,24}$ as n_s increases probably merely indicates that the spectral radius is less accurately approximated by $\bar{\rho}_{1,24}$ for larger values of n_s than for smaller values.[1]

Table 4 presents the averaged per-cycle convergence factors, $\bar{\rho}_{1,12}$, for relaxation alone for Henyey-Greenstein scattering ($q = 2$ in Eq. (10)). Here, in contrast to the results in Table 3, we see that for small- and moderate-sized grids, relaxation performs very well. In fact, for the smallest grids, where $n_\theta = 32$, relaxation performs even better than indicated, as the iterations reduce the error to the level of machine precision in fewer than 12 iterations. Although $\bar{\rho}_{1,12}$ increases with n_θ, the convergence factors are so small that we expect it to be difficult to make improvements using a multigrid method for realistic problem sizes. (Since our discretization is spectrally accurate in n_θ, very large values of n_θ are probably irrelevant.) Experiments and further discussion regarding this point are given in Section 9.3.

---

[1]This is a common effect in linear iteration methods. To understand how it arises, note that in effect, we are applying the power method to determine the spectral radius of the iteration matrix. How fast the power method converges depends on how many eigenvalues there are with modulus close, but not equal, to the spectral radius. When the mesh is larger, there are more such eigenvalues.

	$n_\theta = 32$	$n_\theta = 64$	$n_\theta = 128$	$n_\theta = 256$	$n_\theta = 512$
$n_s = 32$	0.057	0.057	0.115	0.254	0.467
$n_s = 64$	0.059	0.059	0.111	0.245	0.448
$n_s = 128$	0.061	0.061	0.107	0.238	0.433
$n_s = 256$	0.062	0.062	0.104	0.231	0.420
$n_s = 512$	0.064	0.065	0.101	0.224	0.407

Table 4: *Average convergence factor per iteration in first 12 relaxation iterations for test problem 2, with the Henyey-Greenstein scattering operator. Discretization uses second-order upstream differencing in space.*

8   The angular multigrid method

Since the convergence of the relaxation method proposed in Section 7 deteriorates as n_θ increases, it is natural to attempt to accelerate the method using a coarse-grid correction. As Eq. (1) is elliptic in the θ-direction, but advective in the x- and y-directions, we use relaxation sweeps in the (x, y)-plane, as described in Section 7, as the error smoothing procedure, and coarsen only in the θ-direction [6]. In Section 9, we will report results for standard *V*- and *W*-cycles [6]. For simplicity, we assume here that n_θ is a power of 2, and that the coarsest grid has the 4 direction vectors $(\pm 1, 0)$ and $(0, \pm 1)$. We now describe the various components of the coarse-grid correction process in detail.

8.1   Transfer of residuals from fine to coarse grids

In this subsection, we consider two angular grids. The points of the finer angular grid are denoted, as in Section 7, by

$$\Theta_l = l\Delta\Theta, \quad -n_\Theta/2 + 1 \leq l \leq n_\Theta/2 \,,$$

and those on the coarser angular grid are

$$2l\Delta\Theta, \quad -n_\Theta/4 + 1 \leq l \leq n_\Theta/4 \,.$$

(Note that our assumptions, stated earlier, that n_θ is a power of 2 and that the coarsest grid has 4 direction vectors imply that $n_\Theta/4$ is an integer.)

We denote by Ω_F°, $\partial\Omega_F^{out}$, and $\partial\Omega_F^{in}$ the interior and the outflow and inflow boundaries of the finer grid; similarly, Ω_C°, $\partial\Omega_C^{out}$, and $\partial\Omega_C^{in}$ are the interior and the outflow and inflow boundaries of the coarser grid. Residuals are calculated in all interior and outflow boundary points of the fine grid. To describe how the residuals are transferred to interior and outflow boundary points of the coarse grid, the following notation is useful. For integers p and n with $n \geq 2$, n even,

$$p \bmod n = p + zn \,, \quad z = \text{integer chosen so that } -n/2 + 1 \leq p \bmod n \leq n/2 \,. \tag{38}$$

With this notation, the transfer of residuals to the coarse grid is described by

$$R_{j,k,l}^C = \frac{r_{j,k,(2l-1) \bmod n_\Theta}^F + 2r_{j,k,2l}^F + r_{j,k,(2l+1) \bmod n_\Theta}^F}{4} \,, \tag{39}$$

where r^F denotes the residual on the fine grid, and R^C denotes the right-hand side of the correction equation on the coarse grid, as in Eq. (33).

To ensure that the residuals that appear in the expression on the right-hand side of Eq. (39) are well-defined, we must convince ourselves that the points $(j\Delta s, k\Delta s, 2l\Delta\Theta)$ and $(j\Delta s, k\Delta s, ((2l\pm 1)\bmod n_\Theta)\Delta\Theta)$ do not belong to the inflow boundary $\partial\Omega_F^{in}$. Indeed, our definition of the discrete outflow boundary in Section 4.1 implies immediately that for $(j\Delta s, k\Delta s) \in \partial\Omega$ and $-n_\Theta/4 + 1 \le l \le n_\Theta/4$,

$$ (j\Delta s, k\Delta s, l\cdot 2\Delta\Theta) \in \partial\Omega_C^{out} \quad\Rightarrow $$

$$ (j\Delta s, k\Delta s, 2l\Delta\Theta) \in \partial\Omega_F^{out} \quad\text{and}\quad (j\Delta s, k\Delta s, ((2l\pm 1)\bmod n_\Theta)\Delta\Theta) \in \partial\Omega_F^{out} \,. \tag{40} $$

Our definition of the discrete inflow and outflow boundaries is designed to make (40) valid and, thereby, allow transfer of residuals from fine grids to coarse grids using the standard formula (39). We note that (40) would *not* be valid with the first, seemingly more natural definition of *inflow* and *outflow* directions in Section 4.1. This is why we modified the definition in Section 4.1.

For later reference, we note that (40) can also be written as follows:

$$ \left(j\Delta s, k\Delta s, \frac{l}{2}\cdot 2\Delta\Theta\right) \in \partial\Omega_C^{out} \quad\Rightarrow\quad (j\Delta s, k\Delta s, l\Delta\Theta) \in \partial\Omega_F^{out} $$

$$ \text{for } -n_\Theta/2 + 1 \le l \le n_\Theta/2,\ l \text{ even,} \tag{41} $$

and

$$ \left(j\Delta s, k\Delta s, \left(\frac{l\pm 1}{2}\bmod n_\Theta/2\right)\cdot 2\Delta\Theta\right) \in \partial\Omega_C^{out} \quad\Rightarrow\quad (j\Delta s, k\Delta s, l\Delta\Theta) \in \partial\Omega_F^{out} $$

$$ \text{for } -n_\Theta/2 + 1 \le l \le n_\Theta/2,\ l \text{ odd.} \tag{42} $$

8.2 Transfer of corrections from coarse to fine grids

Corrections are interpolated from coarse to fine grids by piecewise linear interpolation. To write down what this means explicitly, we again consider the two angular grids of Section 8.1, and assume that an approximation

$$ f^F_{j,k,l}, \quad 0 \le j,k \le n_s, \quad -n_\Theta/2 + 1 \le l \le n_\Theta/2 $$

to the solution on the fine grid is given, and a correction

$$ f^C_{j,k,l}, \quad 0 \le j,k \le n_s, \quad -n_\Theta/4 + 1 \le l \le n_\Theta/4 $$

has been computed on the coarse grid. The new approximation on the fine grid is then

$$ f^F_{j,k,l} + \left\{ \begin{array}{ll} f^C_{j,k,l/2} & \text{if } l \text{ is even} \\[2ex] \dfrac{f^C_{j,k,(l+1)/2 \bmod n_\Theta/2} + f^C_{j,k,(l-1)/2 \bmod n_\Theta/2}}{2} & \text{if } l \text{ is odd} \end{array} \right\}. \tag{43} $$

If $(j\Delta s, k\Delta s, l\Delta\Theta)$ is an inflow boundary point on the fine grid, then the expression in curly brackets in Eq. (43) ought to be zero: Inflow boundary values should not be altered by the coarse-grid correction. Indeed this is so, since

$$(j\Delta s, k\Delta s, l\Delta\Theta) \in \partial\Omega_F^{in} \quad\Rightarrow\quad \left(j\Delta s, k\Delta s, \frac{l}{2} \cdot 2\Delta\Theta\right) \in \partial\Omega_C^{in}$$

$$\text{for } -n_\Theta/2 + 1 \leq l \leq n_\Theta/2, \, l \text{ even,} \qquad (44)$$

and

$$(j\Delta s, k\Delta s, l\Delta\Theta) \in \partial\Omega_F^{in} \quad\Rightarrow\quad \left(j\Delta s, k\Delta s, \left(\frac{l \pm 1}{2} \bmod n_\Theta/2\right) \cdot 2\Delta\Theta\right) \in \partial\Omega_C^{in}$$

$$\text{for } -n_\Theta/2 + 1 \leq l \leq n_\Theta/2, \, l \text{ odd,} \qquad (45)$$

In fact, (44) is the contrapositive of (41), and (45) is the contrapositive of (42); thus (44) and (45), taken together, are equivalent to (40).

8.3 Solution of the problem on the coarsest grid

On the coarsest grid ($n_\Theta = 4$), we have found that downstream Gauss-Seidel iteration is always a very fast solver. To reduce the residual by a factor of 10^{-10}, we never need more than 4 iterations.

9 Numerical Results

9.1 Multigrid convergence for screened Rutherford scattering

In this subsection, we present numerical results for the equation with the Flatland analog of screened Rutherford scattering: $q = 3$ in Eq. (10). We find convergence to be fast for all grids if D is not too large (Section 9.1.3); convergence does deteriorate eventually as D gets large (see Section 9.1.4), but not nearly as quickly as in the angular domain decomposition method [3]. For a given value of D, the mean free path $\bar{\lambda}$ appears to have little effect on the convergence speed (Section 9.1.5).

9.1.1 Choosing optimal multigrid parameters

Our first experiments are aimed at choosing the parameters of the multigrid method (V- or W-cycles, the number and ordering of pre- and post-relaxation sweeps) in order to achieve the most efficient performance possible. In these experiments, we fix $D = 0.1$, $\bar{\lambda} = 1/20,000$, and consider only the second-order upstream discretization of Test Problem 2.

To compare the effectiveness of different multigrid schemes, one must take into account the amount of work required for a multigrid cycle. For instance, a *V*-cycle with μ pre-relaxation sweeps (relaxation sweeps before coarse-grid correction) and ν post-relaxation

sweeps (relaxation sweeps following coarse-grid correction) on each level requires approximately the equivalent of $(\mu + \nu)(1 + 1/2 + 1/4 + ...) = 2(\mu + \nu)$ relaxation sweeps on the finest level. We approximate the cost of the fine-to-coarse and coarse-to-fine transfers as the equivalent of another 2 relaxation sweeps on the finest level, so that the approximate total cost of a V-cycle is the equivalent of

$$w = 2(\mu + \nu + 1)$$

relaxation sweeps on the finest grid. A similar calculation for W-cycles yields

$$w = \log_2(n_\theta)(\nu + \mu + 1).$$

We call $\left(\bar{\rho}_{M,N}\right)^{1/w}$ (see Eq. 35 for the definition of $\bar{\rho}_{M,N}$) the *effective convergence factor* during iterations M through N. Because a multigrid cycle costs the equivalent of approximately w relaxation sweeps on the finest grid, effective multigrid convergence factors can be compared directly with the convergence factors in Table 3, for instance.

  We called the ordering defined by Eq. (34) the "red-black" ordering of the grid points. The "black" points of a given grid belong to the next coarser grid, whereas the "red" ones don't. We could equally well use the "black-red" ordering:

$$l = -n_\Theta/2 + 2, -n_\Theta/2 + 4, ..., n_\Theta/2, \quad -n_\Theta/2 + 1, -n_\Theta/2 + 3, ..., n_\Theta/2 - 1.$$

Since the "black" points belong to the coarser grid, it is plausible that it should be most efficient to end the pre-relaxation in the "red" points, and to begin the post-relaxation in those points. Numerical experiments confirm this heuristic reasoning. We therefore always use the black-red ordering for pre-relaxation, and the red-black ordering for post-relaxation.

  Table 5 shows effective convergence factors for V-cycles, using $(\mu, \nu) = (1, 0)$, $(0, 1)$, $(1, 1)$, and $(2, 1)$, respectively. The $(0, 1)$-cycles are the most efficient, although they are only slightly more efficient than the other choices. Most notably, all choices of (μ, ν) lead to scalable performance, *i.e.*, effective convergence factors that don't degrade as n_s and n_θ increase. Table 6 shows the effective convergence factors for W-cycles, using $(\mu, \nu) = (0, 1)$. Here, we see some degradation as n_θ increases. This is because, although the values of $\bar{\rho}_{1,24}$ are comparable for V- and W-cycles, the dependence of w on $\log_2(n_\theta)$ in the W-cycle case leads to degradation in the effective convergence factors.

9.1.2 Asymptotic convergence

To illustrate *asymptotic* convergence speed, we reproduce a portion of Table 5 presenting $\left(\bar{\rho}_{11,60}\right)^{1/w}$. Thus we perform here a large number (60) of multigrid cycles, and do not include the first 10 cycles in the average to avoid initialization-dependent transient effects. If we did this with test problems 1 or 2, we would reach machine precision long before completing the 60 cycles and, thus, the results would be polluted by round-off errors. Therefore, we consider test problem 3 here, starting with the initial guess specified in Eq. (37). The results are given in Table 7. They do not differ much from the corresponding results of Table 5. One does see a bit of a deterioration of convergence factors as n_s increases in Table 7; we have not been able to explain that. Notice that such an effect is also seen, although quite slight, in Table 5.

		$n_\theta = 32$	$n_\theta = 64$	$n_\theta = 128$	$n_\theta = 256$	$n_\theta = 512$
$(\mu,\nu) = (1,0)$ $w = 4$ $\left(\bar{\rho}_{1,24}\right)^{1/w}$ is shown	$n_s = 32$	0.699	0.729	0.714	0.707	0.715
	$n_s = 64$	0.701	0.736	0.741	0.706	0.713
	$n_s = 128$	0.704	0.737	0.752	0.727	0.711
	$n_s = 256$	0.706	0.735	0.754	0.749	0.711
	$n_s = 512$	0.709	0.732	0.757	0.756	0.734
$(\mu,\nu) = (0,1)$ $w = 4$ $\left(\bar{\rho}_{1,24}\right)^{1/w}$ is shown	$n_s = 32$	0.698	0.702	0.701	0.706	0.714
	$n_s = 64$	0.701	0.707	0.705	0.706	0.712
	$n_s = 128$	0.703	0.708	0.714	0.706	0.710
	$n_s = 256$	0.706	0.708	0.715	0.708	0.711
	$n_s = 512$	0.708	0.709	0.726	0.712	0.711
$(\mu,\nu) = (1,1)$ $w = 6$ $\left(\bar{\rho}_{1,12}\right)^{1/w}$ is shown	$n_s = 32$	0.716	0.760	0.748	0.745	0.747
	$n_s = 64$	0.712	0.763	0.768	0.754	0.755
	$n_s = 128$	0.709	0.761	0.777	0.759	0.757
	$n_s = 256$	0.705	0.757	0.776	0.774	0.758
	$n_s = 512$	0.702	0.754	0.773	0.779	0.761
$(\mu,\nu) = (2,1)$ $w = 8$ $\left(\bar{\rho}_{1,8}\right)^{1/w}$ is shown	$n_s = 32$	0.663	0.764	0.792	0.782	0.781
	$n_s = 64$	0.659	0.765	0.809	0.792	0.790
	$n_s = 128$	0.655	0.762	0.810	0.805	0.799
	$n_s = 256$	0.652	0.759	0.807	0.815	0.804
	$n_s = 512$	0.648	0.755	0.803	0.815	0.809

Table 5: *Effective convergence factors for V-cycles. Test problem 2, discretized with second-order upstream differencing.*

		$n_\theta = 32$	$n_\theta = 64$	$n_\theta = 128$	$n_\theta = 256$	$n_\theta = 512$
$(\mu,\nu) = (0,1)$ $w = 2\log_2(n_\theta)$ $\left(\bar{\rho}_{1,24}\right)^{1/w}$ is shown	$n_s = 32$	0.866	0.888	0.904	0.917	0.928
	$n_s = 64$	0.868	0.890	0.905	0.917	0.927
	$n_s = 128$	0.869	0.891	0.908	0.917	0.927
	$n_s = 256$	0.870	0.891	0.908	0.917	0.927
	$n_s = 512$	0.871	0.892	0.908	0.918	0.927

Table 6: *Effective convergence factors for W-cycles. Test problem 2, discretized with second-order upstream differencing.*

9.1.3 Dependence of multigrid convergence on the discretization

Here, we again fix the values $D = 0.1$ and $\bar{\lambda} = 1/20,000$. Effects of varying D and $\bar{\lambda}$ will be explored in Sections 9.1.4 and 9.1.5. Table 8 shows $\bar{\rho}_{1,24}$ for test problem 2, discretized with first-order upstream differencing, for various different values of n_s and n_θ.

Multigrid convergence is slightly slower for second-order differencing than for first-order differencing, but in both cases convergence is rapid for all mesh widths.

		$n_\theta = 32$	$n_\theta = 64$	$n_\theta = 128$	$n_\theta = 256$	$n_\theta = 512$
$(\mu,\nu) = (0,1)$ $w = 4$ $\left(\bar{\rho}_{11,60}\right)^{1/w}$ is shown	$n_s = 32$	0.694	0.717	0.702	0.640	0.636
	$n_s = 64$	0.694	0.719	0.725	0.683	0.632
	$n_s = 128$	0.693	0.729	0.735	0.711	0.647
	$n_s = 256$	0.692	0.750	0.767	0.728	0.683
	$n_s = 512$	0.691	0.763	0.811	0.771	0.726

Table 7: *A portion of Table 5 reproduced using $\left(\bar{\rho}_{11,60}\right)^{1/w}$ as the convergence measure, for test problem 3.*

		$n_\theta = 32$	$n_\theta = 64$	$n_\theta = 128$	$n_\theta = 256$	$n_\theta = 512$
$(\mu,\nu) = (0,1)$ $w = 4$ $\left(\bar{\rho}_{1,24}\right)^{1/w}$ is shown	$n_s = 32$	0.695	0.696	0.700	0.707	0.714
	$n_s = 64$	0.696	0.698	0.701	0.706	0.712
	$n_s = 128$	0.700	0.700	0.702	0.704	0.710
	$n_s = 256$	0.701	0.705	0.703	0.705	0.709
	$n_s = 512$	0.706	0.706	0.706	0.705	0.708

Table 8: *A portion of Table 5 reproduced with first-order instead of second-order discretization.*

9.1.4 Dependence of multigrid convergence on the angular diffusion coefficient

As D becomes large, the convergence of the angular multigrid method deteriorates eventually. However, convergence is still excellent for moderately large values of D for which the domain decomposition method described in [3] does not work well. Table 9 illustrates this. The table shows $\left(\bar{\rho}_{1,24}\right)^{1/4}$ for $(0,1)$-V-cycles for the second-order discretization of test problem 2 on a $128 \times 128 \times 128$-grid, with $\bar{\lambda} = 1/20,000$ and various values of D.

D	0.05	0.1	0.2	0.5	1	2	10	100
$\left(\bar{\rho}_{1,24}\right)^{1/4}$	0.716	0.714	0.706	0.705	0.713	0.728	0.837	0.947

Table 9: *Average effective convergence factor per $(0,1)$-V-cycle in first 24 iterations for test problem 2, discretized using second-order upstream differencing, on a $128 \times 128 \times 128$-grid, as a function of D.*

The increase in the convergence factors seen for this discretization as D increases may not be surprising. As $D \to \infty$, the discrete system becomes, to leading order, a set of uncoupled diffusion equations in θ with periodic boundary conditions. Multigrid convergence is well-known to suffer for one-dimensional diffusion with periodic boundary conditions, unless an additional projection step is added to handle the null-space. The coupling provided by the spatial advection terms does, of course, remain crucial in the limit of the continuum equations as $D \to \infty$; they cause the equation to change its nature to spatial diffusion in this limit (see Section 2.4).

9.1.5 Dependence of multigrid convergence on the mean free path

Table 10 illustrates the behavior of the multigrid convergence factors in the Fokker-Planck limit as $\bar{\lambda} \to 0$, with $D = 0.1$ fixed. As $\bar{\lambda}$ decreases (with the parameter ε in Eq. (10) adjusted so that $D = 0.1$ remains fixed), the convergence deteriorates slightly, but multigrid convergence in the limit is only slightly worse than for $\bar{\lambda} = 1/10,000$ (or even $\bar{\lambda} = 1/10$).

$\bar{\lambda}$	10^{-1}	10^{-2}	10^{-3}	10^{-4}	10^{-5}	10^{-6}	10^{-7}	10^{-8}	10^{-9}
$\left(\bar{\rho}_{1,24}\right)^{1/4}$	0.704	0.704	0.706	0.713	0.716	0.717	0.717	0.717	0.717

Table 10: *Average effective convergence factor per $(0,1)$-V-cycle in first 24 iterations for test problem 2, discretized using second-order upstream differencing, on a $128 \times 128 \times 128$-grid, as a function of $\bar{\lambda}$.*

We note that the discretization of the Fokker-Planck operator $\partial^2/\partial\theta^2$ obtained in the limit as $\bar{\lambda} \to 0$ is spectral. If, instead, the standard second-order finite-difference quotient is used, multigrid convergence factors are almost precisely the same as those shown in Table 10; see Table 11.

9.2 Multigrid convergence for the Fokker-Planck equation

Table 11 reproduces a portion of Table 5, but with the right-hand side of the equation replaced by the Fokker-Planck operator, discretized using the three-point stencil. Convergence is nearly identical to Table 5.

		$n_\theta = 32$	$n_\theta = 64$	$n_\theta = 128$	$n_\theta = 256$	$n_\theta = 512$
	$n_s = 32$	0.699	0.699	0.702	0.709	0.718
$(\mu, \nu) = (0, 1)$	$n_s = 64$	0.702	0.702	0.703	0.708	0.716
$w = 4$	$n_s = 128$	0.704	0.704	0.705	0.708	0.714
$\left(\bar{\rho}_{1,24}\right)^{1/w}$ is shown	$n_s = 256$	0.705	0.707	0.707	0.708	0.713
	$n_s = 512$	0.708	0.709	0.724	0.709	0.713

Table 11: *A portion of Table 5 reproduced with the three-point discretization of the Fokker-Planck operator in place of the discretized scattering operator.*

9.3 Multigrid convergence for Henyey-Greenstein scattering

Table 12 presents results similar to some of the results of Table 5, but now the right-hand side of the equation is the Henyey-Greenstein scattering operator. Because of the extraordinarily fast convergence of the multigrid cycles for this scattering kernel, we only show the convergence factor averaged over the first twelve iterations, as this is often sufficient to achieve accuracy near the level of machine precision.

These results should be compared with those of Table 4. For Henyey-Greenstein scattering, it is quite difficult for the coarse-grid correction to improve efficiency compared to

		$n_\theta = 32$	$n_\theta = 64$	$n_\theta = 128$	$n_\theta = 256$	$n_\theta = 512$
	$n_s = 32$	0.489	0.497	0.547	0.570	0.588
$(\mu,\nu) = (0,1)$	$n_s = 64$	0.493	0.496	0.543	0.567	0.586
$w = 4$	$n_s = 128$	0.496	0.497	0.539	0.563	0.582
$\left(\bar\rho_{1,12}\right)^{1/w}$ is shown	$n_s = 256$	0.500	0.500	0.535	0.559	0.578
	$n_s = 512$	0.503	0.505	0.531	0.555	0.574

Table 12: *Results similar to those of Table 5 with Henyey-Greenstein instead of screened Rutherford scattering.*

the very fast convergence of relaxation sweeps alone. For $n_\theta = 1024$ and $n_s = 32$, continued degradation in the convergence factors of relaxation alone is seen (with $\bar\rho_{1,12} = 0.793$), while stabilization is seen in the effective convergence factor for the multigrid V(0,1) cycle (with $\left(\bar\rho_{1,12}\right)^{1/w} = 0.582$); however, it must be noted that such grid sizes are unlikely to be relevant given the spectral accuracy of the discretization in n_θ. Also note that for Henyey-Greenstein scattering, the multigrid convergence has not yet reached an asymptotic convergence range with convergence factors independent of grid size even for $n_\theta = 512$, although numerical experiments show a stabilization for larger grid sizes.

10   Discussion

We have presented a successful two-dimensional extension of the idea of angular multigrid iteration for transport problems [17]. It remains to be investigated whether a method along similar lines can be effective for real, three-dimensional transport problem; if there is a fundamental obstacle, it is not apparent to us.

The algorithm that we use is different from that of [18] in several ways, and which of the differences is responsible for the difference in performance is not, at this point, entirely clear. We intend to investigate this question in the near future, but offer some thoughts about the differences between our algorithm and that of [18] here.

First, we consider a simpler problem than that of [18]. For reasons yet to be understood, we do not encounter, in our problem, the ray effects that were identified as the principal obstacle to efficient performance in [18]. We view our discretization as the analogue of an S_N-discretization in three dimensions, since it uses collocation in angle, so we would expect ray effects to play a role; however, it seems that they do not.

When we generalize our approach to three dimensions in future work, the resulting algorithm will be quite different from that of [18]. Our guiding principle is to first design an efficient algorithm for the Fokker-Planck equation, then create a modification of it that applies to the Boltzmann transport equation. This point of view suggests, in particular, smoothing procedures different from that used in [18].

The definition of the discrete inflow and outflow boundaries is centrally important to the success of our method, as the key implication (40) follows from it. When we began the work on this project, we were less careful about defining the discrete inflow and outflow boundaries; an efficient multigrid scheme could still be constructed, but we had to treat

the transfers between grids at the boundaries with great care to avoid a deterioration in convergence factors with increasing n_s. The approach presented here is simpler, and also seems more likely to generalize to non-rectangular geometries.

Local mesh refinement in x and y could probably be introduced quite easily into our scheme. Local refinement in the θ-direction, especially in an (x, y)-dependent manner, will be more difficult.

For the Henyey-Greenstein scattering kernel, multigrid acceleration does not appear useful, due to the very effective performance of the proposed relaxation scheme itself in this case. Intuitively, we believe that the reason why relaxation works so well for Henyey-Greenstein scattering is that there is less angular diffusion than for screened Rutherford scattering with the same value of D; compare for instance Figures 14 and 15 of [3].

We have assumed in this paper that all scattering is forward-peaked. A more realistic model, known as the Boltzmann-Fokker-Planck equation [7, 15], allows rare "catastrophic" collisions resulting in large-angle deflections. In our Flatland model problem, a reasonable first approximation would be to replace the probability density p by

$$p_{BFP} = \frac{\gamma}{2\pi} + (1 - \gamma)p, \tag{46}$$

where γ is small but positive. Eq. (46) reflects the assumption that large-angle scattering is isotropic. Preliminary numerical results show that for values of γ corresponding to two or three large-angle collisions per 10,000 small-angle collisions, the performance of our multigrid algorithm deteriorates only very slightly. In future work, we will investigate Boltzmann-Fokker-Planck models more extensively, in particular abandoning the simplifying assumption that large-angle scattering is isotropic.

While the assumption of mono-energetic transport is a reasonable first approximation for electron beams, it is of course not physical. As we expand our work to three dimensions, we will also include energy-dependence in future work. Typical discretizations in energy are based on dividing the particles into discrete mono-energetic groups and considering the scattering between groups [8, 16]. As long as particles only lose energy in collisions, the discretized equations accounting for these losses can be viewed as block lower-triangular when organized by energy and can be solved by a forward sweep ordered from the highest energy to the lowest. In the presence of both up- and down-scattering in energy, sweeps downward in energy may still be used as a block Gauss- Seidel iteration.

It is interesting to compare the performance of our algorithm with that of the domain decomposition algorithm of [3]. At first sight, the convergence factors shown in Tables II and III of [3] seem much better than those in the present paper. However, one must of course take into account the cost per iteration for a correct comparison. For the Fokker-Planck equation, this cost is $O(n_s^2 n_\theta)$ for both methods, and therefore the angular domain decomposition method of [3] is in fact much faster than our angular multigrid method. However, for the transport equation, the computational cost per iteration is proportional to the number of nonzero entries in the matrix representing the discretized problem, $O(n_s^2 n_\theta^2)$, for our method, whereas it is $O(n_s^2 n_\theta^3)$ for the method of [3]. The reason is that the sweeps in the "time-like" directions in [3] require the solution of systems of dense linear equations of size $\sim n_\theta \times n_\theta$. For large n_θ, our method is therefore much more efficient than that of [3]. Neither our multigrid approach nor the domain decomposition approach of [3] have

yet been generalized to three dimensions. However, the advantage of the angular multigrid method over the angular domain decomposition method for high-resolution calculations is expected to become more pronounced in three dimensions.

Acknowledgments

References

[1] E. A. Abbott.
Flatland: a romance of many dimensions.
Seely & Co, 1884.

[2] E. Boman, J. Tervo, and M. Vauhkonen.
Modelling the transport of ionizing radiation using the finite element method.
Phys. Med. Biol., 50:265–280, 2005.

[3] C. Börgers.
A fast iterative method for computing particle beams penetrating matter.
J. Comput. Phys., 133:323–339, 1997.

[4] C. Börgers.
Complexity of Monte Carlo and deterministic dose-calculation methods.
Phys. Med. Biol., 43:517–528, 1998.

[5] C. Börgers and E. W. Larsen.
On the accuracy of the Fokker-Planck and Fermi pencil beam equations for charged particle transport.
Medical Physics, 23:1749–1759, 1996.

[6] W. L. Briggs, V. E. Henson, and S. F. McCormick.
A Multigrid Tutorial.
SIAM Books, Philadelphia, 2000.
Second edition.

[7] M. Caro and J. Ligou.
Treatment of scattering anisotropy of neutrons through the Boltzmann-Fokker-Planck equation.
Nucl. Sci. Eng, 83:242–252, 1983.

[8] J. Duderstadt and W. Martin.
Transport Theory.
Wiley, New York, 1979.

[9] K. A. Gifford, J. L. Horton Jr., T. A. Wareing, G. Failla, and F. Mourtada.
Comparison of a finite-element multigroup discrete-ordinates code with Monte Carlo for radiotherapy calculations.

Phys. Med. Biol., 51:2253–2265, 2006.

[10] G. J. Habetler and B. J. Matkowsky.
Uniform asymptotic expansions in transport theory with small mean free paths, and the diffusion approximation.
J. Math. Phys., 16:846–854, 1975.

[11] L. G. Henyey and J. L. Greenstein.
Diffuse radiation in the galaxy.
Astrophys. J., 93:70–83, 1941.

[12] K. R. Hogstrom and P. R. Almond.
Review of electron beam therapy physics.
Phys. Med. Biol., 51:R455–R489, 2006.

[13] S. C. Klevenhagen.
Physics and Dosimetry of Therapy Electron Beams.
Medical Physics Publishing, Madison, WI, 1993.

[14] E. W. Larsen and J. B. Keller.
Asymptotic solution of neutron transport problems for small mean free paths.
J. Math. Phys., 15:75–81, 1974.

[15] E. W. Larsen and L. Liang.
The atomic mix approximation for charged particle transport.
SIAM J. Appl. Math, 68(1):43–58, 2007.

[16] E. W. Larsen and J. E. Morel.
Advances in discrete-ordinates methodology.
In Yoursy Y. Azmy and Enrico Sartori, editors, *Nuclear Computational Science: A Century in Review*. Kluwer Academic Publishing, in press.

[17] J. E. Morel and T. A. Manteuffel.
An angular multigrid acceleration technique for S_n equations with highly forward peaked scattering.
Nuclear Science and Engineering, 107:330–342, 1991.

[18] S. D. Pautz, J. E. Morel, and M. L. Adams.
An angular multigrid acceleration method for S_n equations with highly forward-peaked scattering.
In *Proceedings of the International Conference on Mathematics and Computation, Reactor Physics and Environmental Analyses in Nuclear Applications*, volume I, pages 647–656. Senda Editorial, 1999.

[19] G. C. Pomraning.
The Fokker-Planck operator as an asymptotic limit.
Math. Mod. Meth. Appl. Sci., 2(1):21–36, 1992.

[20] L. Reimer.
Scanning Electron Microscopy.
Springer-Verlag, Berlin, 1985.

[21] S. Reyntjens and R. Puers.
A review of focused ion beam applications in microsystem technology.
J. Micromech. Microeng., 11:287–300, 2001.

[22] D. Schulz-Ertner and H. Tsujii.
Particle radiation therapy using proton and heavier ion beams.
J. Clin. Oncol., 25(8):953–964, 2007.

[23] K. Stüben and U. Trottenberg.
Multigrid methods: Fundamental algorithms, model problem analysis and applications.
In *Multigrid Methods, Lecture Notes in Mathematics*, volume 960, pages 1–176. Springer-
Verlag, 1982.

[24] S Vandewalle and G Horton.
Fourier mode analysis of the multigrid waveform relaxation and time-parallel multigrid meth-
ods.
Computing, 54(4):317–330, 1995.

[25] C. D. Zerby and F. L. Keller.
Electron transport theory, calculations, and experiments.
Nucl. Sci. Eng., 27:190–218, 1967.

[26] J. F. Ziegler, J. P. Biersack, and J. Littnart.
The Stopping Power and Range of Ions in Solids.
Pergamon, New York, 1985.

Figures
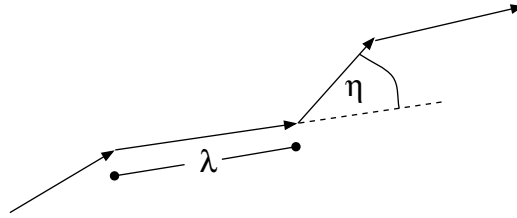
Figure 1: An example of a particle path.
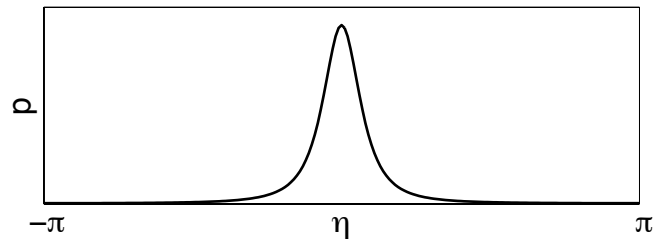
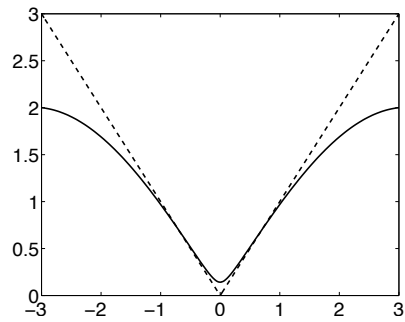Figure 2: A qualitative plot of the probability density p.

Figure 3: $|\eta|$ (dashes) vs. $\sqrt{1 - \cos\eta + \varepsilon^2}$ (solid curve) with $\varepsilon = 0.1$.
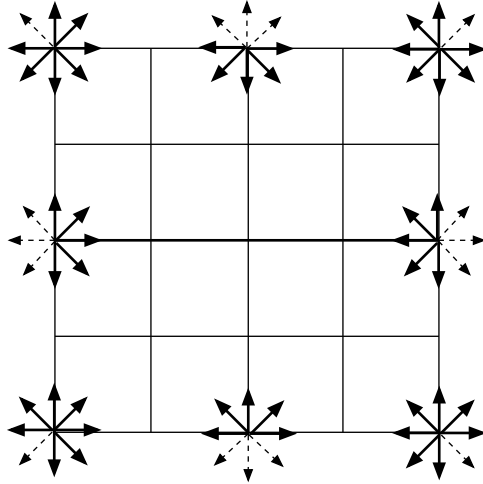
Figure 4: "Outflow", $\partial\Omega^{out}$, (dashes) vs. "inflow", $\partial\Omega^{in}$, (bold) directions on the boundary of Ω.
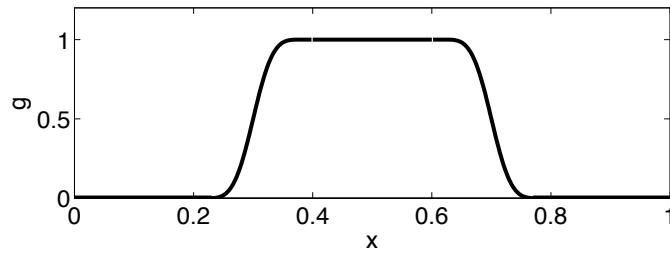

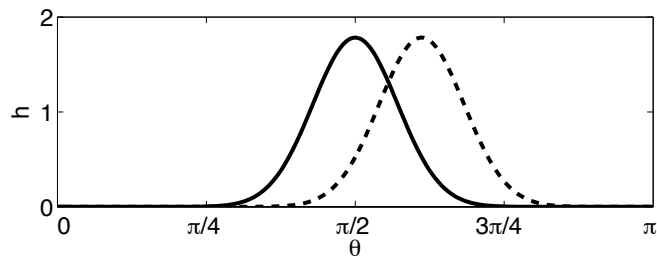Figure 5: Spatial profile of the beams in test problems 1 and 2 at the inflow boundary.

Figure 6: Angular profile of the beams in test problems 1 and 2 (solid and dashed) at the inflow boundary.
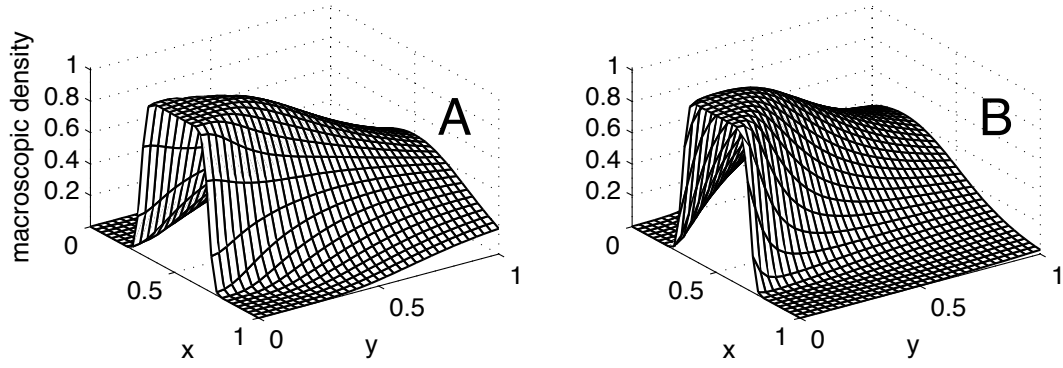
Figure 7: Macroscopic density for two broad beams, one (A) aligned with the grid, and the other (B) not.