SIAM J. SCI. COMPUT.
Vol. 35, No. 5, pp. S250–S276

LOCAL FOURIER ANALYSIS OF SPACE-TIME RELAXATION AND MULTIGRID SCHEMES*

S. FRIEDHOFF†, S. MACLACHLAN†, AND C. BÖRGERS†

Abstract. We consider numerical methods for generalized diffusion equations that are motivated by the transport problems arising in electron beam radiation therapy planning. While Monte Carlo methods are typically used for simulations of the forward-peaked scattering behavior of electron beams, rough calculations suggest that grid-based discretizations can provide more efficient simulations if the discretizations can be made sufficiently accurate, and optimal solvers can be found for the resulting linear systems. The multigrid method for model two-dimensional transport problems presented in [C. Börgers and S. MacLachlan, *J. Comput. Phys.*, 229 (2010), pp. 2914–2931] shows the necessary optimal scaling with some dependence on the choice of scattering kernel. In order to understand this behavior, local Fourier analysis can be applied to the two-grid cycle. Using this approach, expressions for the error-propagation operators of the coarse-grid correction and relaxation steps, projected onto the fine-grid harmonic spaces, can be found. In this paper, we consider easier problems of the form of generalized diffusion problems in space-time that are analogous to model two-dimensional transport problems. We present local Fourier analysis results for these space-time model problems and compare with convergence factors of Börgers and MacLachlan. Since one of our model problems is the diffusion equation itself, we also compare to convergence factors for the diffusion equation of [S. Vandewalle and G. Horton, *Computing*, 54 (1995), pp. 317–330]. The results presented here show that local Fourier analysis does not offer its usual predictivity of the convergence behavior of the diffusion equation and the generalized diffusion equations until we consider unrealistically long time intervals.

Key words. electron transport, diffusion, space-time multigrid, local Fourier analysis

AMS subject classifications. 65F10, 65M22, 65M55

DOI. 10.1137/120881361

1. Introduction.

Beams of charged particles penetrating matter are of interest in many fields, such as cancer therapy using electrons [12, 13], electron microscopy [19], and interstellar radiation theory [11]. Typically, these beams are modeled by Monte Carlo simulations. However, a rough theoretical complexity estimate presented in [5] suggests that deterministic, grid-based methods can provide more efficient simulations if the discretizations can be made sufficiently accurate and if optimal solvers can be found for the resulting linear systems. Depending on the type of application, different parameter choices are used to model the problem mathematically. In this paper, we focus on parameter choices motivated by the electron-beam cancer therapy dose-calculation problem introduced in [12, 13], which implies forward-peaked scattering.

Mathematically, a particle beam, or a configuration of several such beams, is modeled by a boundary value problem for a Boltzmann equation [16]. Applied to electron-beam cancer therapy planning, the equation is assumed to be linear,

$$\frac{1}{c}\frac{\partial f}{\partial t} + \mathbf{v}\cdot\nabla f + \sigma_t(\mathbf{x},E)f - \mathcal{K}_s f = 0, \tag{1.1a}$$

$$f(\mathbf{x},\mathbf{v},E,t_0) = f_0(\mathbf{x},\mathbf{v},E), \tag{1.1b}$$

$$f(\mathbf{x},\mathbf{v},E,t) = g(\mathbf{x},\mathbf{v},E,t) \text{ for } \mathbf{x}\in\partial\Omega, \tag{1.1c}$$

*Received by the editors June 15, 2012; accepted for publication (in revised form) May 1, 2013; published electronically October 28, 2013. This work was supported by the NSF under grant DMS-1015370.

http://www.siam.org/journals/sisc/35-5/88136.html

†Department of Mathematics, Tufts University, Medford, MA 02155 (stephanie.friedhoff@tufts.edu, scott.maclachlan@tufts.edu, christoph.borgers@tufts.edu).

where $\mathbf{x} \in \Omega$ for a given domain, Ω, denotes the particle position, unit vector $\mathbf{v}$ is the particle direction, E is energy, t denotes time, and $f = f(\mathbf{x}, \mathbf{v}, E, t)$ is the phase space density, i. e., the number of particles per unit $(\mathbf{x}, \mathbf{v})$-volume. Furthermore, c is the particle speed, σ_t is the probability of interaction per unit of distance traveled with $\sigma_t = (\overline{\lambda})^{-1}$ for mean free path $\overline{\lambda} > 0$, and $\mathcal{K}_s$ is the scattering kernel. Thus, equation (1.1a) describes how the phase space density evolves in time. In particular, the term $\mathbf{v} \cdot \nabla f$ represents the movement of particles, and $\sigma_t(\mathbf{x}, E)f - \mathcal{K}_s f$ describes the scattering, i. e., particles collide with probability σ_t and get a new direction according to $\mathcal{K}_s$. In electron-beam cancer therapy, the main interest is in time-independent problems, as the transient phase is quite short compared to the total exposure time. Thus, in the computational solution of the linear Boltzmann equation, there are six independent variables: position and velocity in three dimensions (described by direction, $\boldsymbol{v}$, and energy, E).

There are several properties of charged-particle transport which cause difficulties with the accuracy of discretizations and make it hard to find a solution of the discretized linear Boltzmann equation [15, section 3.2]. Therefore, Monte Carlo methods are typically used for simulations of the forward-peaked scattering behavior of electron beams in radiation therapy. More precisely, the Monte Carlo technique simulates the random trajectories of individual particles by using their probability distributions [20]. Grid-based discretizations of the linear Boltzmann equation, (1.1a), however, may provide more efficient simulations if optimal solvers can be found for the resulting linear systems [5]. In the case of mono-energetic linear particle transport in two space dimensions, such algorithms are proposed in [4] and [3]. For high-resolution calculations, the angular multigrid method from [3] is substantially more efficient than the angular domain decomposition method from [4].

In order to extend these results to three-dimensional transport, a better understanding of the results from [3] is necessary. In this paper, we perform local Fourier analysis (LFA), also called local mode analysis, of the multigrid and relaxation methods whose numerical results are presented in [3]. LFA was originally proposed by Brandt [6] and allows estimated predictions of the convergence behavior of multigrid methods when exact predictions by means of rigorous Fourier analysis (model problem analysis) are no longer possible [8, 21, 22]. For many applications, LFA is predictive, i. e., theoretical convergence factors correspond excellently to experimentally measured ones [23]. Numerical results presented in this paper, however, will show that for the class of generalized diffusion problems, LFA does not offer its usual predictivity until we consider unrealistically long time intervals. This result is consistent with those observed in the substantial literature concerning the use of block circulant with circulant block (BCCB) preconditioners for block Toeplitz with Toeplitz block (BTTB) matrices [9]. In particular, typical theoretical results for BCCB preconditioners for BTTB matrices apply only asymptotically, bounding the number of eigenvalues away from unity only as the matrix size goes to infinity [9, 14], even in the non-Hermitian case relevant here. We note that this result does not imply that LFA is "wrong" as a theory, just that it is, in essence, an asymptotic theory and that we are not interested in the asymptotic regime for these problems. As "Flatland" transport is a far too simplified model, extension of the algorithms from [3] (or [4]) to full three-dimensional transport is critical for these approaches to lead to practical benefit. The analysis presented here represents a first step to that extension. A second step, underway in [10], extends this analysis to a setting that is more directly applicable to three-dimensional transport.

This paper is organized as follows. First, in section 2, we review the model problem of [4] describing steady-state mono-energetic linear transport in two space

dimensions and the angular multigrid method of [3] for this model problem, including experimentally measured convergence factors. In section 3, we introduce generalized diffusion problems as a simple subclass of the model problem and describe different space-time relaxation schemes and a two-grid method that corresponds to a two-grid version of the angular multigrid method considered in section 2. Section 4 starts with a review of rigorous, and local, Fourier analysis followed by details of considering LFA to analyze the performance of the space-time relaxation schemes and the two-grid method. Finally, in section 5, we present Fourier analysis results for the space-time relaxation schemes and two-grid method applied to space-time model problems, followed by a discussion in section 6.

2. Transport in Flatland. In the following, we will review the model problem of [4] and the angular multigrid method of [3] for this model problem.

2.1. The model problem. We consider mono-energetic linear particle transport in two space dimensions. More precisely, we consider particle transport in "Flatland" [1], i. e., in a fictitious two-dimensional world, which is simpler than the projection of three-dimensional particle transport into a plane. We consider the motion of particles in a domain, $\Omega \subseteq \mathbb{R}^2$, and assume that all particles move at the same constant speed $c > 0$. The path of each particle results from collisions with the background at random times, causing random directional changes. Since all particles move at the same constant speed, we can characterize the path by two quantities, the distance between two collisions, $\lambda > 0$, and the change of direction, η, caused by each collision; see Figure 2.1.

We assume that the intercollision distances are independent of each other and exponentially distributed with expectation $\overline{\lambda} > 0$, called the mean free path. Directional changes caused by collisions are characterized by their scattering angle, η. We also assume that the scattering angles are independent of each other and of λ with probability density function $p : (-\pi, \pi] \to \mathbb{R}$. Since scattering is assumed to be forward-peaked, $p(\eta)$ should be a continuous function that attains its maximum around $\eta = 0$. Furthermore, particles should have no preferences for scattering to the right over scattering to the left or vice versa; thus, $p(\eta)$ should be an even function. Taking into account all these properties, the graph of p qualitatively looks like a sharply peaked Gaussian.

The model problem describing steady-state mono-energetic linear particle transport in Flatland, applied to electron-beam radiation therapy, is a special case of the continuous linear Boltzmann transport equation, (1.1a). Considering two space dimensions only, we have

$$\mathbf{v} = (\cos(\theta),\ \sin(\theta))^T \quad \text{for } -\pi < \theta \leq \pi.$$

Furthermore, we assume that all interactions result in scattering with probability

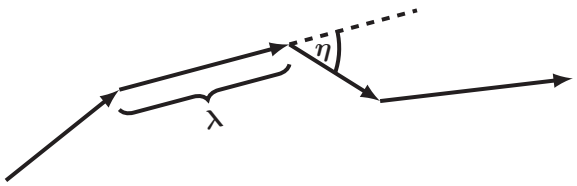


FIG. 2.1. *Sketch of a particle path.*

density $p(\eta)$,

$$[\mathcal{K}_s f](\theta) = \frac{1}{\overline{\overline{\lambda}}} p * f = \frac{1}{\overline{\overline{\lambda}}} \int_{-\pi}^{\pi} p(\eta) f(\mathbf{x}, \theta - \eta)\, d\eta.$$

Note that by constructing the model scattering kernel $\mathcal{K}_s$ with the convolution kernel p derived above, we have a scattering kernel that is qualitatively similar to realistic three-dimensional scattering kernels for charged-particle transport.

With our simplifying assumptions, the continuous Boltzmann transport equation becomes

$$\cos(\theta) f_x + \sin(\theta) f_y = Qf, \tag{2.1a}$$

where the collision operator, Q, is defined by

$$Qf = \frac{1}{\overline{\overline{\lambda}}}(p * f - f) = \frac{1}{\overline{\overline{\lambda}}} \left[\int_{-\pi}^{\pi} p(\eta) f(\theta - \eta)\, d\eta - f(\theta) \right]. \tag{2.1b}$$

We supplement the transport equation (2.1a) with inflow boundary conditions. That is, on parts of the boundary where the direction given by $(\cos(\theta), \sin(\theta))^T$ points into the domain, the value of $f(\boldsymbol{x}, \theta)$ is prescribed.

2.2. Discretization. As derived in [4], the idea of angular discretization is to think of the collision operator Q on the right-hand side of the transport equation (2.1a) coming from an integral scattering term defined by the Fourier coefficients of p. Expressing the scattering operator, Qf, in terms of the Fourier coefficients of p and f, it is straightforward to verify that

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

where the Fourier coefficients are defined by

$$\hat{p}_n = \int_{-\pi}^{\pi} \exp(in\eta) p(\eta)\, d\eta \quad \text{and} \quad \hat{f}_n = \frac{1}{2\pi} \int_{-\pi}^{\pi} \exp(in\tau) f(\tau)\, d\tau.$$

We choose a positive integer, n_θ, divisible by 2 for simplicity, and define $\Delta\theta = (2\pi)/n_\theta$. The discretization of Q is an operator, $Q^{\Delta\theta}$, that maps the space of 2π-periodic functions defined on the grid

$$\Gamma_{n_\theta} = \left\{ \theta_l = l\Delta\theta \ : \ -\frac{n_\theta}{2} + 1 \le l \le \frac{n_\theta}{2} \right\}$$

into itself. The result of angular discretization is

$$(Q^{\Delta\theta} \boldsymbol{f})_l = \frac{1}{n_\theta} \sum_{n=-\frac{n_\theta}{2}+1}^{\frac{n_\theta}{2}} \sum_{m=-\frac{n_\theta}{2}+1}^{\frac{n_\theta}{2}} \frac{\hat{p}_n - 1}{\overline{\lambda}} \cos(n(\theta_l - \theta_m)) f(\theta_m). \tag{2.2}$$

This form of the equation shows that $Q^{\Delta\theta}$ is circulant for any choice of $p(\eta)$.

For the spatial discretization, we use first- or second-order upstream differencing, that is, in the direction opposite to the direction vector $(\cos(\theta_l), \sin(\theta_l))^T$, based on

a uniform two-dimensional spatial grid with mesh size $\Delta s = \frac{1}{n_s}$. These methods are based on the one-sided difference formulas

$$\varphi'(s) = \frac{\varphi(s) - \varphi(s - \Delta s)}{\Delta s} + \mathcal{O}(\Delta s) \text{ or}$$

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

For each of f_x and f_y, we choose the signs of Δs (independently for the x- and y-derivatives) based on the signs of $(\cos(\theta_l), \sin(\theta_l))^T$.

2.3. The Fokker–Planck limit. The Fokker–Planck limit describes the case that collisions become infinitely frequent and infinitesimally weak, i. e., forward-peaked in such a way that the two effects balance each other. In this limit, the linear Boltzmann equation becomes the Fokker–Planck equation.

Expanding $f(\theta - \eta)$ in a Taylor series around $\eta = 0$ including terms of up to order two only,

$$f(\theta - \eta) \approx f(\theta) - \eta f_\theta(\theta) + \frac{\eta^2}{2} f_{\theta\theta}(\theta),$$

and in the case that collisions are strongly forward-peaked, i. e., $p(\eta)$ is small everywhere except near $\eta = 0$, we have

$$Qf = \frac{1}{\overline{\lambda}}(p * f - f)(\theta) = \frac{1}{\overline{\lambda}} \left[\int_{-\pi}^{\pi} p(\eta) f(\theta - \eta)\, d\eta - f(\theta) \right]$$

$$\approx \frac{1}{\overline{\lambda}} \left[\int_{-\pi}^{\pi} p(\eta) \left(f(\theta) - \eta f_\theta(\theta) + \frac{\eta^2}{2} f_{\theta\theta}(\theta) \right) d\eta - f(\theta) \right] = \kappa f_{\theta\theta}(\theta),$$

where the angular diffusion coefficient, κ, is defined by

$$\kappa = \frac{1}{2\overline{\lambda}} \int_{-\pi}^{\pi} \eta^2 p(\eta)\, d\eta. \tag{2.3}$$

Thus, we expect that Qf is "close" to $\kappa f_{\theta\theta}$ and the transport equation (2.1a) becomes

$$\cos(\theta) f_x + \sin(\theta) f_y = \kappa f_{\theta\theta}, \tag{2.4}$$

the so-called Fokker–Planck equation. This calculation was made rigorous in [4].

For the discretization of the second derivative with respect to θ on the right-hand side of the Fokker–Planck equation (2.4), we use standard second-order three-point finite differences.

2.4. Choice of parameters. A reasonable choice for the probability density, p, that mimics realistic three-dimensional scattering kernels as far as possible in Flatland is

$$p(\eta) = \frac{C}{(2(1 - \cos(\eta)) + \varepsilon^2)^{q/2}} \tag{2.5}$$

with the constant $C > 0$ chosen such that the integral of p is equal to one [4]. Proposition 2 of [4] shows that for this choice of $p(\eta)$, in the limit as $\varepsilon \to 0$, the collisions become infinitesimally weak if and only if $q \geq 1$, and the Fokker–Planck limit is valid if and only if $q \geq 3$. Thus, for $q \geq 3$, the transport equation (2.1a) can

be reasonably approximated by the Fokker–Planck equation (2.4) with κ defined as in (2.3). The borderline case, $q = 3$ in (2.5), defines screened Rutherford scattering in Flatland. In this case, convergence is logarithmic [4]. In three space dimensions, the screened Rutherford scattering cross section [25] has the precisely analogous properties [2, Appendix]. A scattering kernel that does not satisfy the conditions for the Fokker–Planck limit to be valid is the Henyey–Greenstein kernel, which is defined by $q = 2$. In this case, in the limit as $\varepsilon \to 0$, one obtains a pseudodifferential operator proportional to $-\sqrt{-\partial^2/\partial\theta^2}$. In this and some other regards, (2.5) with $q = 2$ resembles the Henyey–Greenstein scattering cross section [11]; see [18, 4].

2.5. The angular multigrid method. Since the transport equation (2.1a) is elliptic in the θ-direction but advective in the x- and y-directions, the angular multigrid method described in [3] combines relaxation sweeps ordered parallel to the (x, y) plane with coarse-grid correction over the θ-direction. The discretization of the transport equation (2.1a) that we wish to solve can briefly be represented as

$$\left(\cos(\theta_l)\partial_x^{\Delta s} + \sin(\theta_l)\partial_y^{\Delta s}\right) f = Q^{\Delta\theta} f. \tag{2.6}$$

Here, $f = (f_{j,k,l})_{0\le j,k\le n_s,\ -n_\theta/2+1\le l\le n_\theta/2}$, the symbols $\partial_x^{\Delta s}$ and $\partial_y^{\Delta s}$ denote upstream discretization of the partial derivatives with respect to x and y, and $Q^{\Delta\theta}$ denotes the spectral discretization of the collision operator Q (see section 2.2) or the three-point discretization of the Fokker–Planck operator $\kappa\partial^2/\partial\theta^2$. Equation (2.6) is supplemented with discrete inflow boundary conditions [3, section 4.1]. Denoting the mesh width of the nth angular grid by $\Delta\Theta = 2^n\Delta\theta$, and the number of angles that belong to this grid by $n_\Theta = n_\theta/2^n$, the points of the nth angular grid are given by

$$\Theta_l = l\Delta\Theta, \quad -\frac{n_\Theta}{2} + 1 \le l \le \frac{n_\Theta}{2}.$$

The auxiliary equations solved on coarser grids are of the form

$$\left(\cos(\Theta_l)\partial_x^{\Delta s} + \sin(\Theta_l)\partial_y^{\Delta s}\right) f = Q^{\Delta\Theta} f + R, \tag{2.7}$$

where R denotes residuals transferred from the next finer grid. Equation (2.7) is supplemented with zero inflow boundary conditions. In the multigrid cycle, we use red-black plane-relaxations 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 or odd-even ordering of the angular grid points in the relaxation scheme:

$$l = -\frac{n_\Theta}{2} + 1, -\frac{n_\Theta}{2} + 3, \ldots, \frac{n_\Theta}{2} - 1, \quad -\frac{n_\Theta}{2} + 2, -\frac{n_\Theta}{2} + 4, \ldots, \frac{n_\Theta}{2}.$$

For each fixed l, we sweep over the spatial grid points in the downstream direction, that is, in the direction of the direction vector $(\cos(\Theta_l), \sin(\Theta_l))^T$. In other words, we associate a plane of (x, y) values with each Θ_l and perform a red-black relaxation sweep over these planes, with downstream ordering within each plane.

Results of the performance of this relaxation method [3, section 7.2] are measured by the factor ρ_ν by which the sum of the squares of residuals is reduced in the νth iteration. The average convergence factor over k successive cycles is simply the geometric mean of $\rho_\nu, \rho_{\nu+1}, \ldots, \rho_{\nu+k-1}$. As our test problem, we consider a broad beam in which the mean direction of incoming particles is not aligned with the grid [3, section 5.2]. The type of scattering is specified by the parameters $q, \overline{\lambda}$, and κ, not

TABLE 1

Average convergence factors per iteration in first 24 *relaxation iterations for the test problem* [3, section 5.2], *screened Rutherford scattering, discretized using second-order upstream differencing in space, reproduced from* [3, Table 3].

	$n_\theta = 32$	$n_\theta = 64$	$n_\theta = 128$	$n_\theta = 256$	$n_\theta = 512$
$n_s = 32$	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 2

Average convergence factors per iteration in first 12 *relaxation iterations for the test problem* [3, section 5.2], *Henyey–Greenstein scattering, discretized using second-order upstream differencing in space, reproduced from* [3, Table 4].

	$n_\theta = 32$	$n_\theta = 64$	$n_\theta = 128$	$n_\theta = 256$	$n_\theta = 512$
$n_s = 32$	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

the less intuitive parameter ε in (2.5). Note that ε can be calculated, using (2.3) and (2.5), from $q, \overline{\lambda}$, and κ. Table 1, taken from [3, Table 3], shows the averaged per-cycle convergence factors for the Flatland analog of screened Rutherford scattering ($q = 3$ in (2.5)) with $\kappa = 0.1$ and $\overline{\lambda} = 1/20,000$, using second-order upstream differencing. For fixed n_s, the measured factors increase significantly with n_θ. For fixed n_θ, however, it appears that the factors can be bounded independently of n_s. Using first-order upstream differencing in space gives similar results.

Table 2, taken from [3, Table 4], presents the averaged per-cycle convergence factors for relaxation alone for Henyey–Greenstein scattering ($q = 2$ in (2.5)). Here, in contrast to the results in Table 1, we see that for small- and moderate-sized grids, relaxation performs very well. However, for $n_\theta = 1024$ and $n_s = 32$, an additional result not included in Table 2, we see continued degradation in the convergence factor to 0.793.

Since the convergence of the above red-black relaxation method deteriorates as n_θ increases, it is natural to attempt to accelerate the method using coarse-grid correction. As mentioned earlier, given the fact that the transport equation (2.1a) is elliptic in the θ-direction and advective in the x- and y-directions, we coarsen only in the θ-direction. This semicoarsening strategy is also consistent with the above numerical results, since these suggest that coarsening in space is not necessary. The intergrid transfer operators must account for the treatment of tangential-to-inflow boundary conditions but are otherwise the standard ones used for semicoarsening. Periodic linear interpolation, full-weighting restriction, and rediscretization to get the coarse-grid operators are used [3, sections 8.1, 8.2].

A rough calculation [3, section 9.1.1] shows that a V(0, 1)-cycle costs about the same as four relaxation sweeps on the finest grid. We call $\overline{\rho}^{1/4}$ for averaged convergence factor $\overline{\rho}$ the effective averaged convergence factor, which can be compared directly with the convergence factors in Table 1, for instance. Table 3 shows effective convergence factors for V(0, 1)-cycles applied to the transport equation where the

TABLE 3

Effective convergence factors, i. e., fourth roots of convergence factors, per iteration in first 24 $V(0,1)$*-cycles for the test problem, screened Rutherford scattering, discretized using second-order upstream differencing, the three-point discretization of the Fokker–Planck operator, and effective convergence factors per iteration in first* 12 $V(0,1)$*-cycles with Henyey–Greenstein scattering, taken from* [3, Tables 5, 11, 12].

		$n_\theta = 32$	$n_\theta = 64$	$n_\theta = 128$	$n_\theta = 256$	$n_\theta = 512$
Screened Rutherford scattering	$n_s = 32$	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
Fokker–Planck operator	$n_s = 32$	0.699	0.699	0.702	0.709	0.718
	$n_s = 64$	0.702	0.702	0.703	0.708	0.716
	$n_s = 128$	0.704	0.704	0.705	0.708	0.714
	$n_s = 256$	0.705	0.707	0.707	0.708	0.713
	$n_s = 512$	0.708	0.709	0.724	0.709	0.713
Henyey–Greenstein scattering	$n_s = 32$	0.489	0.497	0.547	0.570	0.588
	$n_s = 64$	0.493	0.496	0.543	0.567	0.586
	$n_s = 128$	0.496	0.497	0.539	0.563	0.582
	$n_s = 256$	0.500	0.500	0.535	0.559	0.578
	$n_s = 512$	0.503	0.505	0.531	0.555	0.574

right-hand side is the screened Rutherford scattering operator, the Fokker–Planck operator, or the Henyey–Greenstein scattering operator, respectively [3, Tables 5,11,12].

Most notably, we see scalable performance in Table 3, i. e., effective convergence factors that do not degrade as n_s and n_θ increase. Convergence for the Fokker–Planck equation is nearly identical to convergence for the transport equation with screened Rutherford scattering. For Henyey–Greenstein scattering, it is quite difficult for the coarse-grid correction to improve efficency compared to the very fast convergence of relaxation sweeps alone. Considering $n_\theta = 1024$ and $n_s = 32$ (not included in Tables 2 and 3), continued degradation in the convergence factors of relaxation alone is seen, while stabilization is seen in the effective convergence factor for the multigrid V$(0,1)$ cycle (with $\overline{\rho}^{1/4} = 0.582$); however, it must be noted that such grid sizes are unlikely to be relevant given the spectral accuracy of the discretization in n_θ.

In summary, the red-black relaxation scheme as well as the angular multigrid method show good performance with some dependence on the choice of scattering kernel. Considering the extensions of these approaches to three-dimensional transport, it is clear that a deeper understanding of this performance would be beneficial in developing efficient algorithms for the different possible discretizations of the five-dimensional Boltzmann equation that describes transport in three spatial dimensions with two angular dimensions. As is well known in the multigrid literature, LFA often provides simple and accurate predictions of the performance of relaxation and multigrid methods, even when these methods are complicated, such as those discussed here [23]. Thus, we now ask if LFA can help us to understand the behavior of the numerical results discussed above. To answer this question, we consider generalized diffusion problems as a simple subclass of the model two-dimensional transport problems (section 3) and apply LFA to these problems (section 4).

3. Model problems in space-time. Fixing $\theta = \theta_0$ on the left-hand side of the transport equation (2.1a) or the Fokker–Planck equation (2.4), respectively, we get

$$\cos(\theta_0) f_x + \sin(\theta_0) f_y = Qf \qquad \text{or} \qquad \cos(\theta_0) f_x + \sin(\theta_0) f_y = \kappa f_{\theta\theta}.$$

For $\theta_0 = 0$, this reduces to

$$f_x = Qf \qquad \text{or} \qquad f_x = \kappa f_{\theta\theta}. \tag{3.1}$$

It is natural in this case to think of x being time and θ being space. Writing "t" instead of "x" and "x" instead of "θ", in (3.1) we get a generalized diffusion equation for scattering operator Q,

$$u_t = Qu, \tag{3.2}$$

and the diffusion equation itself in the Fokker–Planck limit,

$$u_t = \kappa u_{xx}, \quad \kappa > 0. \tag{3.3}$$

To make the model problems analogous to the transport problems of primary interest, we prescribe u at $t = 0$ and impose the periodicity condition

$$u(x \pm 2\pi) \equiv u(x).$$

3.1. Discretization. We discretize our model problems using a spatial mesh size $\Delta x = 2\pi/n_x$ and a time step size $\Delta t = T/n_t$, where n_x and n_t are positive integers and T denotes the final time. The discretization results in a linear system of the form

$$Au = (I_{n_x} \otimes J_{n_t} - Q_{n_x} \otimes I_{n_t})u = f_{IC}, \tag{3.4}$$

where f_{IC} represents the discretized initial condition, in the unknowns $u_{l,j}$, $l = -\frac{n_x}{2} + 1, \ldots, \frac{n_x}{2}$, and $j = 1, \ldots, n_t$. We use first- (BDF1) or second-order (BDF2) backward differences for the time discretization, leading to

$$J_{n_t} = \frac{1}{\Delta t}\begin{bmatrix} 1 & & & & & \\ -1 & 1 & & & & \\ & -1 & 1 & & & \\ & & -1 & 1 & & \\ & & & \ddots & \ddots & \end{bmatrix} \quad \text{or} \quad J_{n_t} = \frac{1}{\Delta t}\begin{bmatrix} 1.5 & & & & & \\ -2 & 1.5 & & & & \\ 0.5 & -2 & 1.5 & & & \\ & 0.5 & -2 & 1.5 & & \\ & & & \ddots & \ddots & \end{bmatrix}.$$

As in section 2.2, the idea of spatial discretization for (3.2) is to think of Q_{n_x} coming from an integral scattering term defined by the Fourier coefficients of the probability density function p,

$$Q_{n_x} = \begin{bmatrix} q_0 & q_1 & \cdots & q_{n_x/2} & q_{-n_x/2+1} & \cdots & q_{-2} & q_{-1} \\ q_{-1} & q_0 & q_1 & \cdots & q_{n_x/2} & q_{-n_x/2+1} & \cdots & q_{-2} \\ \vdots & \ddots & \ddots & & & \ddots & \ddots & \vdots \\ \vdots & & & & & & & \vdots \\ q_1 & \cdots & q_{n_x/2} & q_{-n_x/2+1} & \cdots & \cdots & q_{-1} & q_0 \end{bmatrix}$$

with

$$q_m = \frac{1}{n_x} \sum_{n=-n_x/2+1}^{n_x/2} \frac{\widehat{p}_n - 1}{\overline{\lambda}} \cos(nm\Delta x); \quad \widehat{p}_n = \int_{-\pi}^{\pi} \mathrm{e}^{-in\eta}\, p(\eta)\, d\eta.$$

In the case of the diffusion equation (3.3), Q_{n_x} comes from periodic central differences,

$$Q_{n_x} = \frac{1}{(\Delta x)^2} \begin{bmatrix} -2\kappa & \kappa & & & \kappa \\ \kappa & -2\kappa & \kappa & & \\ & \ddots & \ddots & \ddots & \\ & & \kappa & -2\kappa & \kappa \\ \kappa & & & \kappa & -2\kappa \end{bmatrix}.$$

In both cases, Q_{n_x} is circulant.

3.2. The two-grid method. We are interested in a two-level algorithm of the multigrid type that combines red-black zebra line-relaxations with lines parallel to the time axis with a semicoarsening strategy using coarsening only in the spatial dimension. The relaxation scheme corresponds to the waveform relaxation discussed by Vandewalle and Horton in [24], and it is analogous to the relaxation used in the angular multigrid method of [3].

To describe our method, we consider two space-time grids, Ω_h and Ω_H. The subscript h represents the pair $(\Delta x, \Delta t)$, where Δx denotes the spatial mesh size and Δt is the time increment. The grid Ω_H is derived from Ω_h by doubling the mesh size in the space dimension only, i. e., H represents the pair $(2\Delta x, \Delta t)$. We assume that grid points are in the order of increasing time and from left to right in space, and we permute A into a red-black block ordering by first considering all unknowns at red grid points, $(j_1\Delta x, j_2\Delta t)$ with j_1 odd, and then all unknowns at black grid points, $(j_1\Delta x, j_2\Delta t)$ with j_1 even. Then, the iteration (error-propagation) operator for the red-black relaxation on the fine grid, Ω_h, can be written in the form

$$S^{\mathrm{RB}} = S^{\mathrm{BLACK}} S^{\mathrm{RED}} = (I - M_B A)(I - M_R A)$$

with

$$A = \begin{bmatrix} A_{RR} & A_{RB} \\ A_{BR} & A_{BB} \end{bmatrix}, \quad M_R = \begin{bmatrix} M_{RR}^{-1} & 0 \\ 0 & 0 \end{bmatrix}, \quad M_B = \begin{bmatrix} 0 & 0 \\ 0 & M_{BB}^{-1} \end{bmatrix}.$$

Our interest is in three red-black schemes:

1. *Red-black Jacobi in space-time*, where

$$M_{RR} = D_{RR} \quad \text{and} \quad M_{BB} = D_{BB}$$

with D_{RR} and D_{BB} denoting the diagonals of A_{RR} and A_{BB}, respectively. A red sweep updates the unknowns associated with each red grid point, assuming that all other values, at all other space-time grid points, are frozen as they were at the beginning of the half sweep. A black sweep updates the black space-time grid points analogously.

2. *Red-black Jacobi in space, Gauss–Seidel in time*, where

$$M_{RR} = D_{RR} - L_{RR}^{\mathrm{time}} \quad \text{and} \quad M_{BB} = D_{BB} - L_{BB}^{\mathrm{time}}$$

with D_{RR} and D_{BB} as above and where $-L_{RR}^{\mathrm{time}}$ and $-L_{BB}^{\mathrm{time}}$ are the strictly lower triangular parts of A_{RR} and A_{BB}, respectively, associated with time connections only. A half sweep over the red space-time grid points updates the unknowns associated with each red grid point, assuming that all values at the same time step are frozen as

they were at the beginning of the half sweep, but values at previous time steps are already updated.

3. *Red-black Gauss–Seidel in space-time*, where

$$M_{RR} = D_{RR} - L_{RR} \quad \text{and} \quad M_{BB} = D_{BB} - L_{BB}$$

with $-L_{RR}$ and $-L_{BB}$ denoting the strictly lower triangular parts of A_{RR} and A_{BB}, respectively. A red sweep updates the unknowns associated with each red grid point, assuming that all values at previous time steps and values associated with red grid points at the same time step and to the left in space are already updated.

We use standard geometric coarse-grid correction in the x-direction, using periodic linear interpolation, P, full-weighting restriction, $R = \frac{1}{2}P^T$, and rediscretization to get the coarse-grid operator, A^c. Our two-level algorithm may be represented by the two-grid iteration matrix,

$$M = S^{\mathrm{RB}}(I - P(A^c)^{-1}RA)S^{\mathrm{RB}}. \tag{3.5}$$

4. Two-grid analysis. We consider solving the linear system of equations (3.4) by the two-grid method described in section 3.2. The convergence of this iterative method can be analyzed by studying the iteration matrix, M, given in (3.5). In [22], rigorous Fourier analysis (model problem analysis) was introduced, which allows exact predictions of the convergence behavior for a certain class of model problems and a certain class of multigrid algorithms. The convergence behavior of problems and methods which do not belong to this class may be estimated by applying LFA, which was introduced by Brandt [6]. The word "local" in LFA refers to the fact that boundary conditions and variations in coefficients are neglected.

In section 4.1, we review the ideas behind rigorous Fourier two-grid analysis and discuss how it can be applied to our model problems described in section 3. Section 4.2 is devoted to a review of the ideas of LFA, and in section 5, we present results of this method applied to our model problems.

4.1. Rigorous Fourier analysis. For simple fine- and coarse-grid operators, relaxation schemes, and intergrid transfer operators, the iteration matrix of the two-grid method, M, becomes block-diagonal when the exponential Fourier basis is used with low-dimensional (in our case, two-dimensional) nonzero blocks along the main diagonal and zeros everywhere else, corresponding to small invariant subspaces of the multigrid algorithm. "Rigorous Fourier analysis" means to analyze M by exploiting this observation. More precisely, rigorous Fourier analysis for two-dimensional problems is based on the fact that block-circulant matrices with circulant blocks can be diagonalized by the discrete set of Fourier modes,

$$\boldsymbol{\varphi}^h(\boldsymbol{\omega_k}) = \left(\varphi^h_{j_1,j_2}(\boldsymbol{\omega_k})\right)_{\substack{j_1=-\frac{n_x}{2}+1,\dots,\frac{n_x}{2},\\ j_2=-\frac{n_t}{2}+1,\dots,\frac{n_t}{2}}}, \quad \varphi^h_{j_1,j_2}(\boldsymbol{\omega_k}) = \mathrm{e}^{-i\omega_{k_x} j_1}\,\mathrm{e}^{-i\omega_{k_t} j_2} \quad (\boldsymbol{\omega_k} \in T_h), \tag{4.1}$$

where

$$T_h = \left\{ \boldsymbol{\omega_k} = \left(\frac{2\pi k_x}{n_x}, \frac{2\pi k_t}{n_t}\right) \;:\; k_x = -\frac{n_x}{2}+1,\dots,\frac{n_x}{2},\ k_t = -\frac{n_t}{2}+1,\dots,\frac{n_t}{2} \right\}.$$

In our case, using a first-order upwind (BDF1) discretization in time, we consider $A = I_{n_x} \otimes J_{n_t} - Q_{n_x} \otimes I_{n_t}$,

$$A = \begin{bmatrix} \frac{1}{\Delta t} I_{n_x} - Q_{n_x} & & & & \\ -\frac{1}{\Delta t} I_{n_x} & \frac{1}{\Delta t} I_{n_x} - Q_{n_x} & & & \\ & -\frac{1}{\Delta t} I_{n_x} & \frac{1}{\Delta t} I_{n_x} - Q_{n_x} & & \\ & & \ddots & \ddots & \\ & & & -\frac{1}{\Delta t} I_{n_x} & \frac{1}{\Delta t} I_{n_x} - Q_{n_x} \end{bmatrix}, \tag{4.2}$$

which is only block-Toeplitz with circulant blocks (BTCB) and not BCCB. However, by adding periodicity in time, i. e., considering a periodic space-time grid, we can make a closely related BCCB operator,

$$A^{\mathrm{per}} = \begin{bmatrix} \frac{1}{\Delta t} I_{n_x} - Q_{n_x} & & & & -\frac{1}{\Delta t} I_{n_x} \\ -\frac{1}{\Delta t} I_{n_x} & \frac{1}{\Delta t} I_{n_x} - Q_{n_x} & & & \\ & -\frac{1}{\Delta t} I_{n_x} & \frac{1}{\Delta t} I_{n_x} - Q_{n_x} & & \\ & & \ddots & \ddots & \\ & & & -\frac{1}{\Delta t} I_{n_x} & \frac{1}{\Delta t} I_{n_x} - Q_{n_x} \end{bmatrix}.$$

This modification turns our temporal differential operator, J_{n_t}, into a circulant operator, $J_{n_t}^{\mathrm{per}}$, where the Fourier ansatz (4.1) is valid. The analogous approach can be used for the second-order upwind (BDF2) discretization in time.

It remains to analyze how relaxation and coarse-grid correction act on the Fourier modes. The required calculations are identical with those for LFA, and we therefore proceed to discussing LFA.

4.2. LFA. As opposed to rigorous Fourier analysis, LFA can be applied to problems involving noncirculant operators such as our original operator, $A = I_{n_x} \otimes J_{n_t} - Q_{n_x} \otimes I_{n_t}$; however, as the change in name implies, it is no longer strictly rigorous [8, 21]. The objective of LFA is to determine the asymptotic convergence behavior ignoring effects of boundaries. All considerations in the context of LFA for the two-dimensional case are based on the infinite grid

$$\Omega_h = \left\{(x,t) = (j_1 \Delta x, j_2 \Delta t) : \ (j_1, j_2) \in \mathbb{Z}^2\right\}$$

and use the fact that any infinite-grid block-Toeplitz matrix with Toeplitz blocks can be diagonalized by a set of continuous Fourier modes,

$$\boldsymbol{\varphi}^h(\boldsymbol{\omega}, x, t) = \mathrm{e}^{-i\frac{\omega_x x}{\Delta x}} \, \mathrm{e}^{-i\frac{\omega_t t}{\Delta t}} \quad \text{for } \boldsymbol{\omega} \in (-\pi, \pi]^2, \ (x,t) \in \Omega_h, \tag{4.3}$$

where each component of $\boldsymbol{\omega}$ varies continuously in the interval $(-\pi, \pi]$. The use of infinite-dimensional spaces and operators gives some technical simplifications in the analysis. However, to get a predictive analysis tool based on the L_2-norm of the infinite-grid operator requires maximizing the LFA symbol over the continuos parameter $\boldsymbol{\omega}$. Thus, in practice, LFA has to be taken from its infinite-grid setting, and we maximize, instead, the symbol over a finite set of values of $\boldsymbol{\omega}$.

The analysis of the two-grid iteration matrix, M, by means of LFA is as follows: We start by extending the discrete operator, A, as well as relaxation and coarse-grid correction operators to the infinite grid, Ω_h, ignoring any boundary conditions. We then use the continuous Fourier modes (4.3) to diagonalize the infinite-grid operator, A^h, corresponding to the discrete operator, A. If A^c is also chosen so that it is a BTTB operator, then it can also be diagonalized, now by coarse-grid Fourier modes, i. e., with ω_x replaced by $2\omega_x$, in the factor-2 coarsening considered here. A discrete set of diagonal terms is then chosen, corresponding to a discrete mesh in $\boldsymbol{\omega}$, needed for the prediction of the performance of the two-grid method. More precisely, we consider the discrete set of the Fourier modes (4.1). Using this Fourier ansatz, (4.2), and properties of the Kronecker product, the LFA predicted eigenvalues of the fine-grid operator, $A = I_{n_x} \otimes J_{n_t} - Q_{n_x} \otimes I_{n_t}$, are

$$\lambda^{(h,\boldsymbol{k})} = \frac{1}{\Delta t}\left(1 - \mathrm{e}^{i\omega_{k_t}}\right) - \sum_{m=-\frac{n_x}{2}+1}^{\frac{n_x}{2}} q_m \, \mathrm{e}^{-i\omega_{k_x} m}$$

if BDF1 is used for discretizing in time, and

$$\lambda^{(h,\boldsymbol{k})} = \frac{1}{\Delta t}\left(\frac{3}{2} - 2\,\mathrm{e}^{i\omega_{k_t}} + \frac{1}{2}\,\mathrm{e}^{i2\omega_{k_t}}\right) - \sum_{m=-\frac{n_x}{2}+1}^{\frac{n_x}{2}} q_m \, \mathrm{e}^{-i\omega_{k_x} m}$$

if we use BDF2 for discretizing in time. On the coarse grid, we have $\omega^H_{k_x} = 2\omega^h_{k_x}$, yielding the LFA predicted eigenvalues of the coarse-grid operator, A^c,

$$\lambda^{(H,\boldsymbol{k})} = \frac{1}{\Delta t}\left(1 - \mathrm{e}^{i\omega_{k_t}}\right) - \sum_{m=-\frac{n_x}{4}+1}^{\frac{n_x}{4}} q^c_m \, \mathrm{e}^{-i2\omega_{k_x} m}$$

for BDF1 in time, or

$$\lambda^{(H,\boldsymbol{k})} = \frac{1}{\Delta t}\left(\frac{3}{2} - 2\,\mathrm{e}^{i\omega_{k_t}} + \frac{1}{2}\,\mathrm{e}^{i2\omega_{k_t}}\right) - \sum_{m=-\frac{n_x}{4}+1}^{\frac{n_x}{4}} q^c_m \, \mathrm{e}^{-i2\omega_{k_x} m}$$

for BDF2 in time.

For the intergrid transfer operators, R and P, the Fourier modes are no longer eigenfunctions. Since we consider semicoarsening in space, the restriction operator, R, maps two fine-grid functions, the Fourier harmonics, to one coarse-grid function. More precisely, these two functions are associated with the frequencies

$$\boldsymbol{\omega_k} \in \left(-\frac{\pi}{2}, \frac{\pi}{2}\right] \times (-\pi, \pi] \quad \text{and} \quad \boldsymbol{\omega'_k} = \boldsymbol{\omega_k} - \operatorname{sign}(\omega_{k_x}) \begin{bmatrix} \pi \\ 0 \end{bmatrix}$$

(see Figure 4.1). It can be verified that the fine-grid harmonic spaces are left invariant by the coarse-grid correction process [23]. As a result, we can use the continuous Fourier modes to block-diagonalize the infinite-grid restriction and interpolation operators. The size of these blocks is 2×2, reflecting the coupling of the two Fourier harmonics. If $\boldsymbol{r}^h$ is a fine-grid vector composed only of the two harmonic modes $\boldsymbol{\varphi}^h(\boldsymbol{\omega_k})$ and $\boldsymbol{\varphi}^h(\boldsymbol{\omega'_k})$,

$$\boldsymbol{r}^h = c_\omega \boldsymbol{\varphi}^h(\boldsymbol{\omega_k}) + c_{\omega'} \boldsymbol{\varphi}^h(\boldsymbol{\omega'_k}),$$

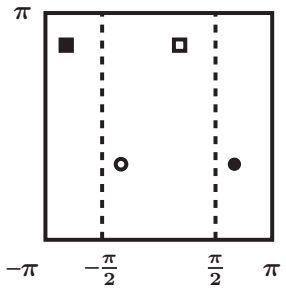


FIG. 4.1. *For a given frequency $\boldsymbol{\omega_k}$ ($\circ$ or $\square$), the frequency $\boldsymbol{\omega'_k}$ for which the corresponding Fourier modes, $\boldsymbol{\varphi}^h(\boldsymbol{\omega_k})$ and $\boldsymbol{\varphi}^h(\boldsymbol{\omega'_k})$, coincide on the coarse grid is marked by $\bullet$ and $\blacksquare$, respectively.*

then, for full-weighting restriction, we obtain

$$R\boldsymbol{r}^h = \left(\cos^2\left(\frac{\omega_{k_x}}{2}\right)c_\omega + \sin^2\left(\frac{\omega_{k_x}}{2}\right)c_{\omega'}\right)\boldsymbol{\varphi}^H(2\omega_{k_x},\omega_{k_t}).$$

Periodic linear interpolation maps the single coarse-grid mode $\boldsymbol{\varphi}^H(2\omega_{k_x},\omega_{k_t})$ to the corresponding two fine-grid harmonics $\boldsymbol{\varphi}^h(\boldsymbol{\omega_k})$ and $\boldsymbol{\varphi}^h(\boldsymbol{\omega'_k})$. More precisely, we obtain

$$P\boldsymbol{\varphi}^H(2\omega_{k_x},\omega_{k_t}) = \cos^2\left(\frac{\omega_{k_x}}{2}\right)\boldsymbol{\varphi}^h(\boldsymbol{\omega_k}) + \sin^2\left(\frac{\omega_{k_x}}{2}\right)\boldsymbol{\varphi}^h(\boldsymbol{\omega'_k}).$$

In order to make a convergence prediction of the discussed two-grid method for our model problems, it remains to analyze the three red-black relaxation schemes. Since we use red-black relaxation, the Fourier modes are again no longer eigenfunctions of the relaxation operators, but the harmonic spaces are left invariant under these operators. As a result, we can use the continuous Fourier modes to block-diagonalize the infinite-grid relaxation operators, and, by multiplying out the 2×2 blocks that we get from relaxation and coarse-grid correction, we can block-diagonalize the two-grid operator, M, as a whole. Then, we choose a discrete set of diagonal blocks, corresponding to a discrete mesh in $\boldsymbol{\omega}$. The asymptotic convergence behavior of the two-grid method can then be predicted by measuring the spectral radius of the transformed two-grid operator over this discrete space. In the following, we discuss how to find analytical expressions for the coefficients of the 2×2 diagonal blocks for our three red-black relaxation schemes. Each of these methods consists of two partial steps. In the first partial step, the red step, only values at red grid points, (j_1, j_2) with j_1 odd, are updated, while values at black grid points, (j_1, j_2) with j_1 even, remain unchanged. The second step, the black step, then only updates values at black grid points. Denoting the partial step relaxation operators by S^{RED} and S^{BLACK}, the complete red-black relaxation operator, S^{RB}, is given by their product, $S^{\mathrm{RB}} = S^{\mathrm{BLACK}}S^{\mathrm{RED}}$. We present primarily results for BDF2 time discretization; however, expressions for BDF1 time discretization are similar.

4.2.1. Red-black Jacobi in space-time. To analyze relaxation, we consider the fine-grid mode $\boldsymbol{\varphi}^h = \boldsymbol{\varphi}^h(\boldsymbol{\omega_k})$, and first express the value of $(S^{\mathrm{RED}}\boldsymbol{\varphi}^h(\boldsymbol{\omega_k}))_{j_1,j_2}$ at each grid point (j_1, j_2). Applying Jacobi in space-time type relaxation to (3.4) of our

discrete model problem, we have

(4.4)
$$\left(S^{\mathrm{RED}}\boldsymbol{\varphi}^h\right)_{j_1,j_2} = \begin{cases} \dfrac{1}{\frac{1}{\Delta t} - q_0}\left[\displaystyle\sum_{m<0} q_m \boldsymbol{\varphi}^h_{j_1+m,j_2} + \sum_{m>0} q_m \boldsymbol{\varphi}^h_{j_1+m,j_2} + \frac{1}{\Delta t}\boldsymbol{\varphi}^h_{j_1,j_2-1}\right] & \text{if } j_1 \text{ odd,} \\ \boldsymbol{\varphi}^h_{j_1,j_2} & \text{if } j_1 \text{ even} \end{cases}$$

in the case of BDF1 time discretization, and

(4.5)
$$\left(S^{\mathrm{RED}}\boldsymbol{\varphi}^h\right)_{j_1,j_2} = \begin{cases} \dfrac{1}{\frac{3}{2\Delta t} - q_0}\left[\displaystyle\sum_{m<0} q_m \boldsymbol{\varphi}^h_{j_1+m,j_2} + \sum_{m>0} q_m \boldsymbol{\varphi}^h_{j_1+m,j_2} + \frac{2}{\Delta t}\boldsymbol{\varphi}^h_{j_1,j_2-1} - \frac{1}{2\Delta t}\boldsymbol{\varphi}^h_{j_1,j_2-2}\right] & \text{if } j_1 \text{ odd,} \\ \boldsymbol{\varphi}^h_{j_1,j_2} & \text{if } j_1 \text{ even} \end{cases}$$

for BDF2 time discretization. We now make the ansatz that $(S^{\mathrm{RED}}\boldsymbol{\varphi}^h)_{j_1,j_2}$ can be written as a linear combination of the two Fourier harmonics,

$$\begin{aligned}(S^{\mathrm{RED}}\boldsymbol{\varphi}^h)_{j_1,j_2} &= c_\omega \,\mathrm{e}^{-i\omega_{k_x} j_1}\,\mathrm{e}^{-i\omega_{k_t} j_2} + c_{\omega'}\,\mathrm{e}^{-i\omega'_{k_x} j_1}\,\mathrm{e}^{-i\omega'_{k_t} j_2} \\ &= \left(c_\omega + c_{\omega'}\,\mathrm{e}^{-i\pi j_1}\right)\mathrm{e}^{-i\omega_{k_x} j_1}\,\mathrm{e}^{-i\omega_{k_t} j_2}\,.\end{aligned}$$

Equating terms, we obtain $c_{\omega'} = 1 - c_\omega$ and

$$c_\omega = \frac{1}{2} + \frac{1}{2\left(\frac{1}{\Delta t} - q_0\right)}\left[\sum_{m<0} q_m\,\mathrm{e}^{-i\omega_{k_x} m} + \sum_{m>0} q_m\,\mathrm{e}^{-i\omega_{k_x} m} + \frac{1}{\Delta t}\,\mathrm{e}^{i\omega_{k_t}}\right]$$

for BDF1 time discretization, and

$$c_\omega = \frac{1}{2} + \frac{1}{\frac{3}{\Delta t} - 2q_0}\left[\sum_{m<0} q_m\,\mathrm{e}^{-i\omega_{k_x} m} + \sum_{m>0} q_m\,\mathrm{e}^{-i\omega_{k_x} m} + \frac{2}{\Delta t}\,\mathrm{e}^{i\omega_{k_t}} - \frac{1}{2\Delta t}\,\mathrm{e}^{i2\omega_{k_t}}\right]$$

for BDF2 time discretization. Analogously, we express the value of $(S^{\mathrm{BLACK}}\boldsymbol{\varphi}^h(\boldsymbol{\omega_k}))_{j_1,j_2}$ at each grid point (j_1, j_2), i. e., equations (4.4) and (4.5) with the expressions for j_1 odd and j_1 even interchanged, and make the ansatz that $(S^{\mathrm{BLACK}}\boldsymbol{\varphi}^h(\boldsymbol{\omega_k}))_{j_1,j_2}$ can be written as a linear combination of the two Fourier harmonics,

$$\begin{aligned}(S^{\mathrm{BLACK}}\boldsymbol{\varphi}^h)_{j_1,j_2} &= d_\omega \,\mathrm{e}^{-i\omega_{k_x} j_1}\,\mathrm{e}^{-i\omega_{k_t} j_2} + d_{\omega'}\,\mathrm{e}^{-i\omega'_{k_x} j_1}\,\mathrm{e}^{-i\omega'_{k_t} j_2} \\ &= \left(d_\omega + d_{\omega'}\,\mathrm{e}^{-i\pi j_1}\right)\mathrm{e}^{-i\omega_{k_x} j_1}\,\mathrm{e}^{-i\omega_{k_t} j_2}\,.\end{aligned}$$

Equating terms again, we obtain $d_\omega = c_\omega$ and $d_{\omega'} = d_\omega - 1$.

4.2.2. Red-black Jacobi in space, Gauss–Seidel in time. Applying Jacobi in space and Gauss–Seidel in time type relaxation to (3.4) of our discrete model problem, using BDF2 discretization in time, we have

(4.6)
$$\left(S^{\mathrm{RED}}\boldsymbol{\varphi}^h\right)_{j_1,j_2} = \begin{cases} \dfrac{1}{\frac{3}{2\Delta t} - q_0}\left[\displaystyle\sum_{m<0} q_m \boldsymbol{\varphi}^h_{j_1+m,j_2} + \sum_{m>0} q_m \boldsymbol{\varphi}^h_{j_1+m,j_2} \right. \\ \qquad\qquad \left. + \dfrac{2}{\Delta t}\left(S^{\mathrm{RED}}\boldsymbol{\varphi}^h\right)_{j_1,j_2-1} - \dfrac{1}{2\Delta t}\left(S^{\mathrm{RED}}\boldsymbol{\varphi}^h\right)_{j_1,j_2-2}\right] & \text{if } j_1 \text{ odd}, \\ \boldsymbol{\varphi}_{j_1,j_2} & \text{if } j_1 \text{ even}. \end{cases}$$

The value of $(S^{\mathrm{BLACK}}\boldsymbol{\varphi}^h(\boldsymbol{\omega_k}))_{j_1,j_2}$ at each grid point (j_1, j_2) is given by (4.6) with the expressions for j_1 odd and j_1 even interchanged. Analogously to section 4.2.1, we make the ansatz

$$\begin{aligned} \left(S^{\mathrm{RED}}\boldsymbol{\varphi}^h\right)_{j_1,j_2} &= c_\omega \,\mathrm{e}^{-i\omega_{k_x} j_1}\,\mathrm{e}^{-i\omega_{k_t} j_2} + c_{\omega'}\,\mathrm{e}^{-i\omega'_{k_x} j_1}\,\mathrm{e}^{-i\omega'_{k_t} j_2}, \\ \left(S^{\mathrm{BLACK}}\boldsymbol{\varphi}^h\right)_{j_1,j_2} &= d_\omega \,\mathrm{e}^{-i\omega_{k_x} j_1}\,\mathrm{e}^{-i\omega_{k_t} j_2} + d_{\omega'}\,\mathrm{e}^{-i\omega'_{k_x} j_1}\,\mathrm{e}^{-i\omega'_{k_t} j_2} \end{aligned}$$

to obtain

$$c_\omega = \frac{1 + \dfrac{1}{\frac{3}{2\Delta t} - q_0}\left[\displaystyle\sum_{m<0} q_m \,\mathrm{e}^{-i\omega_{k_x} m} + \sum_{m>0} q_m \,\mathrm{e}^{-i\omega_{k_x} m} - \frac{2}{\Delta t}\,\mathrm{e}^{i\omega_{k_t}} + \frac{1}{2\Delta t}\,\mathrm{e}^{i2\omega_{k_t}}\right]}{2 - \dfrac{1}{\frac{3}{2\Delta t} - q_0}\left[\dfrac{4}{\Delta t}\,\mathrm{e}^{i\omega_{k_t}} - \dfrac{1}{\Delta t}\,\mathrm{e}^{i2\omega_{k_t}}\right]},$$

$c_{\omega'} = 1 - c_\omega$, $d_\omega = c_\omega$, and $d_{\omega'} = d_\omega - 1$.

4.2.3. Red-black Gauss–Seidel in space-time. Applying Gauss–Seidel in space-time type relaxation to (3.4) of our discrete model problem, using BDF2 discretization in time, we have

(4.7)
$$\left(S^{\mathrm{RED}}\boldsymbol{\varphi}^h\right)_{j_1,j_2} = \begin{cases} \dfrac{1}{\frac{3}{2\Delta t} - q_0}\left[\displaystyle\sum_{m<0} q_m \left(S^{\mathrm{RED}}\boldsymbol{\varphi}^h\right)_{j_1+m,j_2} + \sum_{m>0} q_m \boldsymbol{\varphi}_{j_1+m,j_2} \right. \\ \qquad\qquad \left. + \dfrac{2}{\Delta t}\left(S^{\mathrm{RED}}\boldsymbol{\varphi}^h\right)_{j_1,j_2-1} - \dfrac{1}{2\Delta t}\left(S^{\mathrm{RED}}\boldsymbol{\varphi}^h\right)_{j_1,j_2-2}\right] & \text{if } j_1 \text{ odd}, \\ \boldsymbol{\varphi}_{j_1,j_2} & \text{if } j_1 \text{ even}. \end{cases}$$

The value of $(S^{\mathrm{BLACK}}\boldsymbol{\varphi}^h(\boldsymbol{\omega_k}))_{j_1,j_2}$ at each grid point (j_1, j_2) is given by (4.7) with the expressions for j_1 odd and j_1 even interchanged. Analogously to section 4.2.1, we make the ansatz

$$\begin{aligned} \left(S^{\mathrm{RED}}\boldsymbol{\varphi}^h\right)_{j_1,j_2} &= c_\omega \,\mathrm{e}^{-i\omega_{k_x} j_1}\,\mathrm{e}^{-i\omega_{k_t} j_2} + c_{\omega'}\,\mathrm{e}^{-i\omega'_{k_x} j_1}\,\mathrm{e}^{-i\omega'_{k_t} j_2}, \\ \left(S^{\mathrm{BLACK}}\boldsymbol{\varphi}^h\right)_{j_1,j_2} &= d_\omega \,\mathrm{e}^{-i\omega_{k_x} j_1}\,\mathrm{e}^{-i\omega_{k_t} j_2} + d_{\omega'}\,\mathrm{e}^{-i\omega'_{k_x} j_1}\,\mathrm{e}^{-i\omega'_{k_t} j_2} \end{aligned}$$

to obtain

$$c_\omega = \frac{1 + \frac{1}{\frac{3}{2\Delta t} - q_0}\left[\sum_{m>0} q_m \, \mathrm{e}^{-i\omega_{k_x} m} - \sum_{m<0} q_m \, \mathrm{e}^{-i(\omega_{k_x}+\pi)m} - \frac{2}{\Delta t}\,\mathrm{e}^{i\omega_{k_t}} + \frac{1}{2\Delta t}\,\mathrm{e}^{i2\omega_{k_t}}\right]}{2 - \frac{1}{\frac{3}{2\Delta t} - q_0}\left[\sum_{m<0} q_m \left(\mathrm{e}^{-i\omega_{k_x} m} + \mathrm{e}^{-i(\omega_{k_x}+\pi)m}\right) + \frac{4}{\Delta t}\,\mathrm{e}^{i\omega_{k_t}} - \frac{1}{\Delta t}\,\mathrm{e}^{i2\omega_{k_t}}\right]},$$

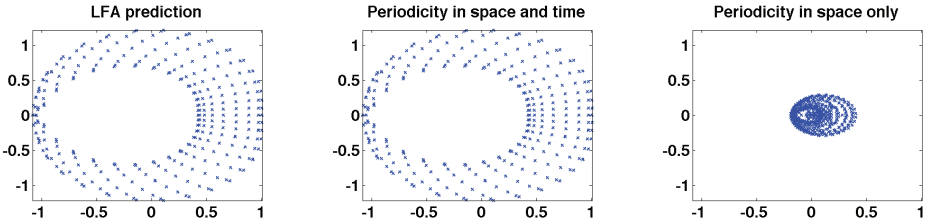
$c_{\omega'} = 1 - c_\omega$, $d_\omega = c_\omega$, and $d_{\omega'} = d_\omega - 1$.

5. Numerical results. In this section, we present Fourier analysis results for our three red-black relaxation schemes, two-grid $(1,1)$-, and two-grid $(0,1)$-cycles. In sections 5.1 and 5.2, we consider the diffusion equation (3.3) with $\kappa = 1$. Because of similarities to the exponential Fourier mode analysis for the multigrid waveform relaxation method discussed by Vandewalle and Horton [24], in section 5.2 we compare our two-grid $(1,1)$-cycle predictions to convergence factors in [24]. Sections 5.3 and 5.4 are devoted to the generalized diffusion equation (3.2).

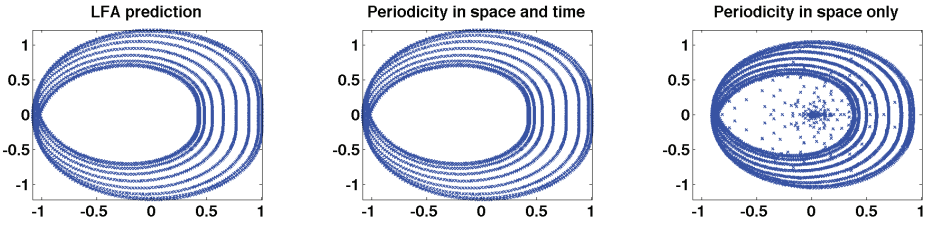
5.1. Fourier analysis for the diffusion equation. We report on tests of eigenvalues of the iteration matrices of the different schemes. For each scheme, we consider two cases, namely, a periodic space-time grid and a space-time grid which is periodic in space only. On the periodic space-time grid, both our differential operators, Q_{n_x} and $J_{n_t}^{\mathrm{per}}$, are circulant and thus diagonizable by the discrete set of Fourier modes. So, we expect LFA to be exact for Jacobi type relaxations, since it is rigorous, and good for Gauss–Seidel type relaxations and the two-grid $(1,1)$-cycles. Considering a grid with periodicity in space only, the spatial differential operator, Q_{n_x}, is still circulant, and thus diagonalizable. However, the temporal differential operator, J_{n_t}, is only Toeplitz, and more precisely, in the case of BDF1 time discretization, a Jordan block, and thus not diagonalizable. Therefore, we expect LFA to be good only for large grids. The two questions that arise naturally are, what does "large" mean, and, is convergence good enough to understand the results of the angular relaxation scheme and the angular multigrid method for transport problems presented in [3]? Since our discretization is spectrally accurate in n_x, and the corresponding spatial differential operator, Q_{n_x}, is circulant, very large values of n_x are probably irrelevant. Thus, in most of our calculations, we have fixed $n_x = 16$, and only vary the number of temporal grid points, n_t, with fixed $\Delta t = 1/16$, by varying T in the definition of n_t, $n_t = T/\Delta t$.

Figure 5.1 shows the eigenvalues in the complex plane for computations on space-time grids of sizes 16×32 ($T = 2$), 16×256 ($T = 16$), and 16×2048 ($T = 128$), respectively, for red-black Jacobi in space-time. The left plots present eigenvalues of the LFA prediction, the middle and right plots show analytically computed eigenvalues, i. e., exact eigenvalues of the iteration matrix. More precisely, the middle plots show results on periodic space-time grids and the right plots results on grids which are periodic in space only. As expected, the LFA prediction matches the analytically computed results on periodic space-time grids exactly, since the Fourier analysis is rigorous on these grids. However, we have to move to a very large number of grid points in the temporal direction to get a close prediction of results on a grid which is periodic in space only. For realistic grid sizes, the LFA prediction does not capture convergence very accurately.

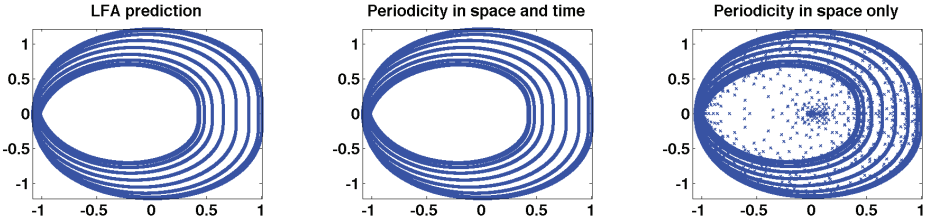
Figures 5.2 and 5.3 present similar results to Figure 5.1 for red-black Gauss–Seidel in space-time relaxation, which, for the diffusion equation, is equivalent to



(a) Eigenvalues on a 16×32 space-time grid



(b) Eigenvalues on a 16×256 space-time grid

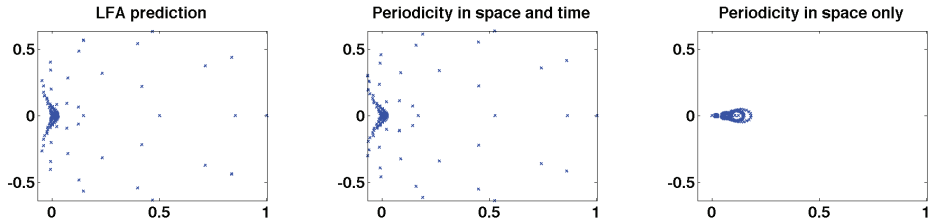


(c) Eigenvalues on a 16×2048 space-time grid

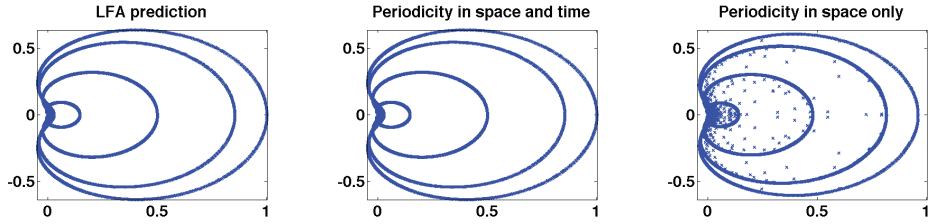
FIG. 5.1. *Eigenvalues on a* (a) 16×32 ($T = 2$), (b) 16×256 ($T = 16$), *and* (c) 16×2048 ($T = 128$) *space-time grid,* $\Delta t = 1/16$ *fixed, red-black Jacobi in space-time, BDF2 time discretization of the diffusion equation with* $\kappa = 1$. *At left, eigenvalues predicted by LFA, in the middle, eigenvalues of the analytically computed iteration matrix on a periodic space-time grid, and at right, analytically computed eigenvalues on a grid which is periodic in space only.*

red-black Jacobi in space, Gauss–Seidel in time relaxation, and two-grid $(1, 1)$-cycles with Gauss–Seidel in space-time relaxation, respectively. Here, in contrast to the results in Figure 5.1, we see that LFA is no longer rigorous for periodic space-time grids. Again, LFA is only asymptotically accurate for results on a grid which is periodic in space only, but not good for a small number of grid points in the temporal direction. Furthermore, LFA has no hope of predicting the behavior of relaxation, since it always predicts relaxation to have a spectral radius of one. Numerical results for the diffusion equation discretized using BDF1 in time show the same limitations.

5.2. Comparison to [24]. Vandewalle and Horton have applied LFA to the discrete-time version of the multigrid waveform relaxation method, presented in [17], for the diffusion equation [24]. This multigrid method uses standard central differences for the spatial operator and BDF1, BDF2, or Crank–Nicolson for the time discretization. Furthermore, the standard smoother is a zebra Gauss–Seidel method, one of the red-black relaxation schemes discussed in section 3.2. Coarse-grid correction is based on semicoarsening, with coarsening in the spatial dimension only, and intergrid transfer operators are linear interpolation and full-weighting restriction in the spatial dimension.

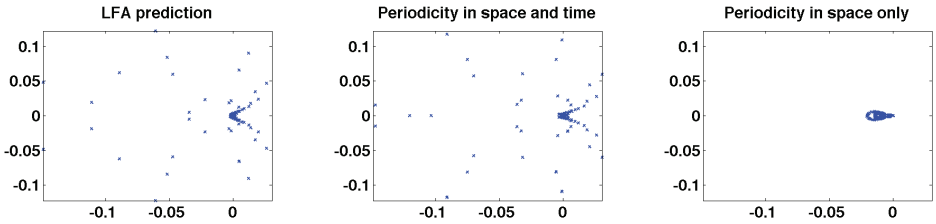


(a) Eigenvalues on a 16×32 space-time grid

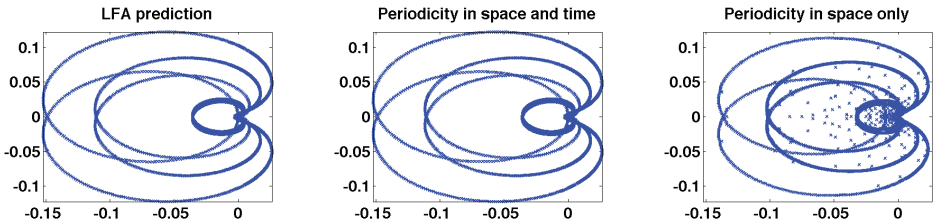


(b) Eigenvalues on a 16×2048 space-time grid

FIG. 5.2. *Results similar to those of Figure* 5.1 *for red-black Gauss–Seidel in space-time instead of red-black Jacobi in space-time relaxation.*



(a) Eigenvalues on a 16×32 space-time grid



(b) Eigenvalues on a 16×2048 space-time grid

FIG. 5.3. *Results similar to those of Figure* 5.1 *for two-grid* $(1,1)$*-cycles with red-black Gauss–Seidel in space-time relaxation instead of red-black Jacobi in space-time relaxation only.*

In [24], Vandewalle and Horton compare numerical results of two-grid $(1,1)$-cycles of the multigrid waveform relaxation method applied to the diffusion equation for different mesh-aspect ratios, $\lambda_h = \Delta t/(\Delta x)^2$, on a grid of size 128×128. They report convergence factors based on the two-norm of the residual, averaged over the first 30 iterations, or, in the case of fast convergence, over iterations up to the point where the initial residual is reduced in norm by a factor of 10^{-10}. Table 4 shows these average convergence factors (taken from [24, Table 1]), $\overline{\rho}$, as well as our computations of the LFA predictions of spectral radii, ρ_{LFA}, and the spectral radii of the analytically

TABLE 4

Average convergence factors per iteration (reported in [24]*),* $\overline{\rho}$*, LFA predictions of spectral radii,* ρ_{LFA}*, and spectral radii of the analytically computed iteration matrix,* $\rho_{analytic}$*, for the diffusion equation with* $\kappa = 1$*, discretized using BDF*1 *or BDF*2 *in time, two-grid* $(1,1)$*-cycles with red-black Gauss–Seidel in space-time relaxation,* 128×128*-grid. Results are truncated after two digits past the decimal point because this corresponds to the precision of the results in* [24].

	$\log_2(\lambda_h)$	-6	-3	0	3	6	9
BDF1	$\overline{\rho}$	0.02	0.11	0.08	0.04	0.01	0.00
	ρ_{LFA}	0.06	0.14	0.09	0.04	0.01	0.00
	ρ_{analytic}	0.00	0.02	0.07	0.04	0.01	0.00
BDF2	$\overline{\rho}$	0.02	0.12	0.12	0.05	0.01	0.00
	ρ_{LFA}	0.07	0.16	0.14	0.06	0.01	0.00
	ρ_{analytic}	0.00	0.01	0.08	0.05	0.01	0.00

computed iteration matrix, ρ_{analytic}, on 128×128 space-time grids that are periodic in space only, given to the same accuracy as in [24]. For these computations of the spectral radii, we have fixed the spatial mesh size, $\Delta x = (2\pi)/128$, and computed the time increment, Δt, from the expression $\Delta t = \lambda_h(\Delta x)^2$ with λ_h given by the value of $\log_2(\lambda_h)$. For the analytic computations, this implicitly varies the final time, T, in the expression $\Delta t = T/128$. For the LFA calculations, we have used enough points in the time-Fourier direction to adequately resolve the LFA spectral radius; this issue is explored more in Table 5.

Table 4 suggests that for small values of λ_h, corresponding to small time increments, Δt, if Δx is fixed, LFA gives neither a good prediction for the numerical convergence behavior nor for the analytical convergence behavior. However, for larger values of λ_h, corresponding to large time increments, Δt, if Δx is fixed, LFA gives a close prediction to both the numerical and analytical convergence behavior. Put another way, for fixed spatial mesh size, $\Delta x = (2\pi)/128$, and fixed time increment, $\Delta t = 1/128$, i. e., considering $T = 1$, we can compute the effective diffusion coefficient, κ, from the expression $\kappa = \lambda_h(\Delta x)^2/\Delta t$, and we can relate the accuracy of the LFA prediction to physical transport parameters using (2.3). From this point of view, small values of λ_h and, thus, small diffusion coefficients, κ, correspond to large mean free paths, $\overline{\lambda}$, in the Fokker–Planck limit, whereas larger values of λ_h correspond to large diffusion coefficients, κ, and, thus, small mean free paths, $\overline{\lambda}$, in the Fokker–Planck limit. Table 4 suggests that for large mean free paths, $\overline{\lambda}$, characterizing "thin" material, LFA gives neither a good prediction for the numerical convergence behavior nor for the analytical convergence behavior. However, for smaller mean free paths, $\overline{\lambda}$, characterizing "thick" material, LFA gives a close prediction to both the numerical and analytical convergence behavior.

In the LFA results presented in Table 4, the number of points in the time-Fourier direction required to attain an accurate LFA prediction depends strongly on λ_h. To see this in more detail, Table 5 presents results for independently varying both the time step and the number of points in the time and the time-Fourier directions. We consider spectral radii of LFA predictions, ρ_{LFA}, and spectral radii of the analytically computed iteration matrix, ρ_{analytic}, for two-grid $(1,1)$-cycles for the diffusion equation with $\kappa = 1$, discretized using BDF2 in time, on $n_x \times n_t$ space-time grids, which are periodic in space only, with fixed $n_x = 16$.

For fixed time increment, Δt, the LFA predictions stabilize if we use $1/\Delta t$ grid points in the time-Fourier direction. The reason for this is that the maximum of the LFA-predicted attenuation factors occurs at a temporal wave number of approximately $2\pi\Delta t$ which is resolved when we have $1/\Delta t$ points in the time-Fourier direction, but

TABLE 5

LFA predictions of spectral radii, ρ_{LFA}, and spectral radii of the analytically computed iteration matrix, $\rho_{analytic}$, for two-grid $(1,1)$-cycles for the diffusion equation with $\kappa = 1$, discretized using BDF2 in time, on $n_x \times n_t$ space-time grids, which are periodic in space only, $n_x = 16$ fixed.

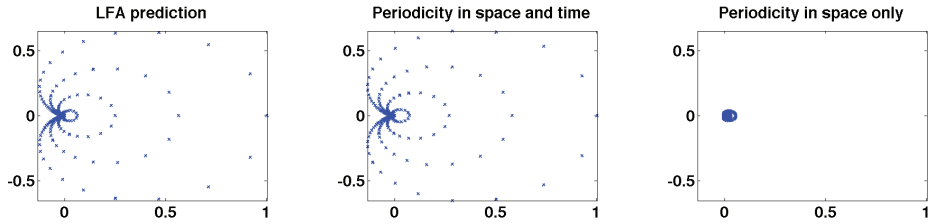
		$\Delta t = 1/16$	$\Delta t = 1/32$	$\Delta t = 1/64$	$\Delta t = 1/128$	$\Delta t = 1/256$
ρ_{LFA}	$n_t = 16$	1.57e-01	1.19e-01	3.78e-02	6.53e-03	8.97e-04
	$n_t = 32$	1.57e-01	1.59e-01	1.27e-01	4.16e-02	7.28e-03
	$n_t = 64$	1.57e-01	1.59e-01	1.59e-01	1.28e-01	4.25e-02
	$n_t = 128$	1.57e-01	1.61e-01	1.59e-01	1.59e-01	1.29e-01
	$n_t = 256$	1.57e-01	1.61e-01	1.62e-01	1.59e-01	1.59e-01
	$n_t = 512$	1.57e-01	1.61e-01	1.62e-01	1.62e-01	1.59e-01
	$n_t = 1024$	1.57e-01	1.62e-01	1.62e-01	1.62e-01	1.62e-01
ρ_{analytic}	$n_t = 16$	1.52e-02	4.18e-03	7.97e-04	1.28e-04	1.81e-05
	$n_t = 32$	2.05e-02	6.12e-03	1.15e-03	1.78e-04	2.52e-05
	$n_t = 64$	3.55e-02	1.18e-02	2.20e-03	3.31e-04	4.72e-05
	$n_t = 128$	5.52e-02	2.69e-02	5.93e-03	8.76e-04	1.21e-04
	$n_t = 256$	8.61e-02	5.02e-02	1.81e-02	2.91e-03	4.01e-04
	$n_t = 512$	1.15e-01	8.69e-02	4.40e-02	1.16e-02	1.59e-03
	$n_t = 1024$	1.35e-01	1.20e-01	8.38e-02	3.78e-02	8.28e-03

not fewer. Furthermore, Table 5 shows that there is a big discrepancy between the analytic spectral radii and the LFA predictions if $n_t \Delta t$ is small. In other words, we need an exceptionally large final time, T, before the LFA and analytic rates are comparable. For example, $\Delta t = 1/16$ and $n_t = 1024$ corresponds to $T = 64$, a time period that is not interesting in practice since it is very long in comparison with the diffusion time scale.

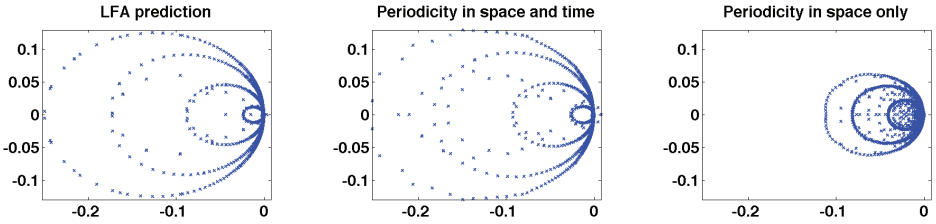
The results in Table 4 show similar behavior as a function of λ_h. For values of λ_h with $\log_2(\lambda_h) > 0$, $\Delta t > 0.02$ and, so, about 50 points in the time-Fourier direction should be sufficient for the LFA results to resolve the maximum attenuation factor predicted. In Table 4, we use 128 points in the time-Fourier direction (for convenience of matching the analytical computation), and see this to be the case. For $\lambda = 1$, $\Delta t = 0.0024$, suggesting about 400 points are needed and, indeed, we find we need at least 256 (for BDF2) or 512 (for BDF1) points to resolve the maximum predicted attenuation factor. For negative $\log_2(\lambda_h)$, $1/\Delta t$ seems to seriously overestimate the number of points needed in the time-Fourier direction; however, these correspond to unrealistically small time steps, so this behavior may not be of great interest. On the other hand, comparing the LFA predictions with the numerical results from [24] or the analytic computations, we see very good matches for λ_h with $\log_2(\lambda_h) > 0$, which correspond to final times $T > 1$. For $\lambda_h = 1$ with 128 temporal grid points, we have $T = 0.3084$ and see good (but not perfect) agreement between the LFA and the computational and analytic results. For negative $\log_2(\lambda_h)$, final times are small, $T < 0.04$, and we see, as expected, poor predictions of true behavior.

5.3. Fourier analysis for model problems in space-time. In this subsection, we examine the dependence of the performance of the red-black relaxation schemes and the two-grid $(0,1)$-cycles for the generalized diffusion equation (3.2) on the choice of scattering kernel. As in section 5.1, we consider the equation discretized on a periodic space-time grid and on a space-time grid which is periodic in space only. Again, fixing $n_x = 16$ and varying the number of grid points in time, n_t, only, numerical experiments show that LFA is not good for small n_t.

Figure 5.4 shows the eigenvalues in the complex plane for computations on a 16×2048 space-time grid for red-black Gauss–Seidel in space-time relaxation and



(a) red-black Gauss-Seidel in space-time relaxation



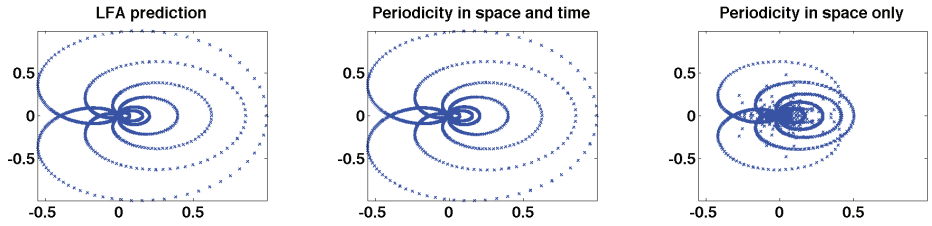
(b) two-grid $(0,1)$-cycles with red-black Gauss-Seidel in space-time relaxation

FIG. 5.4. *Eigenvalues on a* 16×2048 *space-time grid,* (a) *red-black Gauss–Seidel in space-time relaxation, and* (b) *two-grid* $(0,1)$*-cycles with red-black Gauss–Seidel in space-time relaxation, Henyey–Greenstein scattering, BDF2 time discretization. At left, eigenvalues predicted by LFA, in the middle, eigenvalues of the analytically computed iteration matrix on a periodic space-time grid, and at right, analytically computed eigenvalues on a grid which is periodic in space only.*

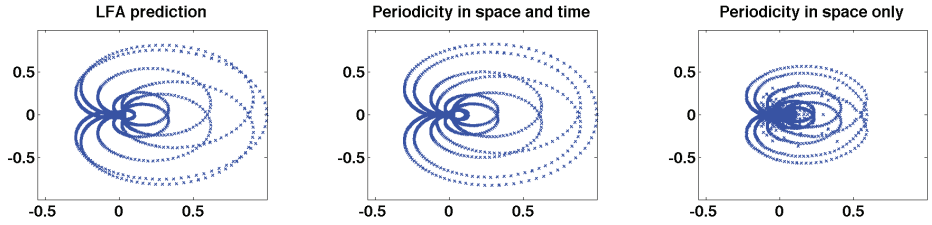
for two-grid $(0,1)$-cycles with red-black Gauss–Seidel in space-time relaxation for the generalized diffusion equation with the Henyey–Greenstein scattering operator ($q = 2$ in (2.5)), discretized using BDF2 in time. We see that, even on this large grid, LFA does not capture the very fast convergence behavior of relaxation, and gives only a poor prediction of two-grid $(0,1)$-cycle results on a space-time grid which is periodic in space only.

As seen in Figures 5.5 and 5.6, LFA gives a better prediction of results for the case of spatial (but not temporal) periodicity for screened Rutherford scattering ($q = 3$ in (2.5)) than it does for Henyey–Greenstein scattering ($q = 2$ in (2.5)). However, again, we need a very large number of grid points in the temporal direction. We also want to note that, due to stronger red-to-black connections, for screened Rutherford scattering, we see that results for red-black Jacobi in space, Gauss–Seidel in time relaxation differ from results for red-black Gauss–Seidel in space-time relaxation.

5.4. Comparison to [3]. Our reason for considering generalized diffusion problems and applying LFA to these problems was to understand the results of the angular multigrid method for transport problems, presented in [3] and reviewed in section 2.5. In this subsection, we present LFA results for two-grid $(0,1)$-cycles for the diffusion equation (3.3) and the generalized diffusion equation (3.2). In section 5.3, we have seen that LFA does not give good performance predictions on coarse time grids in physical space. In this subsection, we analyze the dependence of LFA predictions on the number of points in Fourier space by considering Fourier space-time grids of size $n_x \times (T/\Delta t)$. We want to emphasize that considering $T/\Delta t$ points in the time-Fourier direction is independent of the physical meaning of T as being the total integration time in physical space. For our convenience only, we follow the rule that the number of points in time-Fourier space is given by $T/\Delta t$, so that we can increase this number by considering the two limits, $T \to \infty$ and $\Delta t \to 0$. Effects of these two limits on the



(a) red-black Jacobi in space, Gauss-Seidel in time relaxation



(b) red-black Gauss-Seidel in space-time relaxation

FIG. 5.5. *Results similar to those of Figure* 5.4(a) *with screened Rutherford instead of Henyey–Greenstein scattering,* (a) *red-black Jacobi in space, Gauss–Seidel in time, and* (b) *red-black Gauss–Seidel in space-time relaxation.*

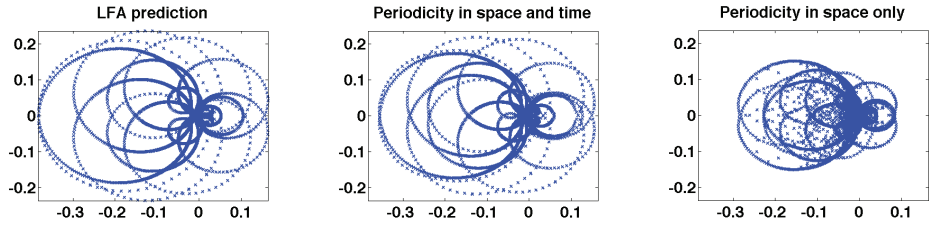


FIG. 5.6. *Results similar to those of Figure* 5.4(b) *with screened Rutherford instead of Henyey–Greenstein scattering.*

LFA prediction of the convergence behavior of two-grid $(0,1)$-cycles with red-black Gauss–Seidel in space-time relaxation for the diffusion equation, the generalized diffusion equation with the screened Rutherford and the Henyey–Greenstein scattering operator are illustrated in Tables 6, 7, and 8. In Tables 6 and 7, it can be seen that the LFA predicts that the performance of two-grid $(0,1)$-cycles for the diffusion equation and for the generalized diffusion equation with the screened Rutherford scattering operator do not change if we go from $T = 1$ to $T = 2$. Changing T has no effect on the equations that are analyzed within the LFA, only on how many points are sampled in the time direction of Fourier space. Thus, observing no changes in the results in Tables 6 and 7 as we increase T, this says that $T = 1$ samples enough points in Fourier space to get an accurate picture of the error-propagation operator.

For the diffusion equation, we see in Table 6 that the LFA results are consistently shifted down by two rows as we move to each subsequent column, i. e., we see consistent results when we double n_x and multiply Δt by $1/4$. This is to be expected as these shifts result in a simple scaling of the (infinite-grid) operator by a factor of four, which will have no effect on the performance of relaxation or coarse-grid correction. For each fixed n_x, the LFA prediction stabilizes for $\Delta t \leq 2/(n_x)^2$, which implies that the mesh-

TABLE 6

LFA predictions of spectral radii for the diffusion equation, discretized using BDF2 in time, two-grid $(0,1)$-cycles with red-black Gauss–Seidel in space-time relaxation, $n_x \times (T/\Delta t)$-grid.

		$n_x = 32$	$n_x = 64$	$n_x = 128$	$n_x = 256$	$n_x = 512$
$T = 1$	$\Delta t = 1/128$	0.247	0.208	0.125	0.058	0.018
	$\Delta t = 1/256$	0.249	0.235	0.167	0.090	0.033
	$\Delta t = 1/512$	0.250	0.247	0.208	0.125	0.058
	$\Delta t = 1/1024$	0.250	0.249	0.235	0.167	0.090
	$\Delta t = 1/2048$	0.250	0.250	0.247	0.208	0.125
	$\Delta t = 1/4096$	0.250	0.250	0.250	0.235	0.167
	$\Delta t = 1/8192$	0.250	0.250	0.250	0.247	0.208
$T = 2$	$\Delta t = 1/128$	0.248	0.208	0.125	0.058	0.018
	$\Delta t = 1/256$	0.249	0.235	0.167	0.090	0.033
	$\Delta t = 1/512$	0.250	0.247	0.208	0.125	0.058
	$\Delta t = 1/1024$	0.250	0.250	0.236	0.167	0.090
	$\Delta t = 1/2048$	0.250	0.250	0.247	0.208	0.125
	$\Delta t = 1/4096$	0.250	0.250	0.250	0.235	0.167
	$\Delta t = 1/8192$	0.250	0.250	0.250	0.247	0.208

TABLE 7

Results similar to those in Table 6 with the generalized diffusion equation with the screened Rutherford scattering operator instead of the diffusion equation.

		$n_x = 32$	$n_x = 64$	$n_x = 128$	$n_x = 256$	$n_x = 512$
$T = 1$	$\Delta t = 1/128$	0.375	0.370	0.356	0.291	0.187
	$\Delta t = 1/256$	0.376	0.370	0.363	0.333	0.248
	$\Delta t = 1/512$	0.376	0.370	0.365	0.351	0.300
	$\Delta t = 1/1024$	0.376	0.370	0.365	0.356	0.331
	$\Delta t = 1/2048$	0.376	0.370	0.365	0.357	0.341
	$\Delta t = 1/4096$	0.376	0.370	0.365	0.357	0.343
	$\Delta t = 1/8192$	0.376	0.370	0.365	0.357	0.343
$T = 2$	$\Delta t = 1/128$	0.375	0.370	0.356	0.291	0.187
	$\Delta t = 1/256$	0.376	0.370	0.364	0.333	0.248
	$\Delta t = 1/512$	0.376	0.370	0.365	0.351	0.300
	$\Delta t = 1/1024$	0.376	0.370	0.365	0.356	0.331
	$\Delta t = 1/2048$	0.376	0.370	0.365	0.357	0.341
	$\Delta t = 1/4096$	0.376	0.370	0.365	0.357	0.343
	$\Delta t = 1/8192$	0.376	0.370	0.365	0.357	0.343

aspect ratio, $\lambda_h = \Delta t/(\Delta x)^2$, is $1/(2\pi^2)$ or smaller. Moreover, for fine-enough time grids in Fourier space, predictions are bounded independently of the time step size if $\lambda_h \leq 1/(2\pi^2)$. To see stabilization in the LFA prediction for $n_x = 512$, for example, we need about $130,000$ points in the time-Fourier direction. Thus, Table 6 suggests that for the diffusion equation we get very slow convergence of the LFA prediction to its asymptotic bound. The mode that gives the maximum predicted attenuation factor occurs at a temporal wave number of approximately $2\pi(n_x^2/128)\Delta t$.

In Table 7, we see that LFA predictions for the generalized diffusion equation with the screened Rutherford scattering operator behave similarly to the LFA predictions for the diffusion equation but with a nicer scaling factor, meaning that we get much faster convergence of the LFA prediction to its asymptotic bound. More precisely, for each fixed n_x, the LFA predictions stabilize for $\Delta t \leq 1/(4n_x)$. We note that, due to the choice of parameters in the scattering operators in [3], the continuum problem to be solved is slightly different for each n_x, leading to slightly different stabilized rates predicted by the LFA.

TABLE 8

Results similar to those in Table 7 with Henyey–Greenstein instead of screened Rutherford scattering.

		$n_x = 32$	$n_x = 64$	$n_x = 128$	$n_x = 256$	$n_x = 512$
$T = 1$	$\Delta t = 1/128$	0.045	0.089	0.161	0.238	0.237
	$\Delta t = 1/256$	0.046	0.089	0.162	0.238	0.237
	$\Delta t = 1/512$	0.046	0.089	0.162	0.238	0.237
$T = 2$	$\Delta t = 1/128$	0.089	0.162	0.238	0.238	0.249
	$\Delta t = 1/256$	0.089	0.162	0.238	0.238	0.249
	$\Delta t = 1/512$	0.089	0.162	0.238	0.238	0.249
$T = 4$	$\Delta t = 1/128$	0.162	0.238	0.238	0.250	0.249
	$\Delta t = 1/256$	0.162	0.238	0.238	0.250	0.249
	$\Delta t = 1/512$	0.162	0.238	0.238	0.250	0.249
$T = 8$	$\Delta t = 1/128$	0.239	0.238	0.250	0.250	0.249
	$\Delta t = 1/256$	0.239	0.238	0.250	0.250	0.249
	$\Delta t = 1/512$	0.239	0.238	0.250	0.250	0.249
$T = 16$	$\Delta t = 1/128$	0.239	0.250	0.250	0.250	0.249
	$\Delta t = 1/256$	0.239	0.250	0.250	0.250	0.249
	$\Delta t = 1/512$	0.239	0.250	0.250	0.250	0.249

Table 8 shows that for Henyey–Greenstein scattering, LFA results differ significantly from LFA results for the diffusion equation and for the generalized diffusion equation with screened Rutherford scattering. This observation is consistent with the observed convergence behavior for the transport equation when comparing Henyey–Greenstein scattering to screened Rutherford scattering and the Fokker–Planck equation. In contrast to convergence predictions for the diffusion equation and for the generalized diffusion equation with the screened Rutherford scattering operator, we see that there is no dependence on Δt. Moreover, the temporal wave number where the maximum LFA-predicted attenuation factor occurs is independent of Δt and T, once T is large enough.

Compared to the effective convergence factors of angular multigrid V$(0,1)$-cycles, given in Table 3, we see that for the diffusion equation theoretical convergence factors correspond excellently to experimentally measured ones: The stabilized LFA prediction for the diffusion equation on a reasonable physical domain is 0.25. Considering unscaled convergence factors of angular multigrid V$(0,1)$-cycles, i. e., the fourth powers of the numbers in Table 3, this prediction is in excellent agreement with the results for the Fokker–Planck equation in Table 3 (the fourth root of 0.25 is 0.71). However, for screened Rutherford and Henyey–Greenstein scattering, the LFA predictions overestimate the observed convergence rates (0.34 and 0.25 for LFA vs. 0.25 and 0.11 for unscaled convergence factors).

6. Discussion. We have presented LFA results for generalized diffusion problems in space-time that are analogous to model two-dimensional transport problems. In contrast to problems that have only elliptic behavior, for parabolic space-time problems, LFA does not offer its usual predictivity of the convergence behavior for relaxation and multigrid schemes unless the time period used for the LFA computations is very long in comparison with the diffusion time scale.

Another key issue in LFA is determining the number of grid points needed in the Fourier domain in order to ensure that we resolve the true LFA prediction of the spectral radii, based on the continuous Fourier-domain representation. For elliptic problems on fixed physical domains (usually the unit interval, square, or cube), it is usually assumed that taking the same number of points in Fourier space as in physical

space is sufficient. Here, we find that for the diffusion equation and the generalized diffusion equation with screened Rutherford scattering, we need approximately as many points in the time-Fourier direction as we need to cover a unit time interval, independent of the actual time interval in question. In contrast, for generalized diffusion equations with Henyey–Greenstein scattering, we need many more points in the time-Fourier direction than are needed to cover a unit time interval. Even with a sufficient number of points in the Fourier domain, the resulting predictions of the convergence behavior of our space-time problems remains inaccurate unless we consider unrealistically long time intervals. We observe that the LFA factors are always greater than the analytical factors, but we have no analysis to prove this observation. As a next step, the ideas of half-space mode analysis [7] can be considered to give more useful predictions [10].

While the LFA gives us some insight into the performance of multigrid methods for the two-dimensional Flatland transport model, we find that it exposes many more questions about the nature and applicability of LFA than it answers. This suggests that it will not, in general, be a useful tool in extending the results from [3] to full three-dimensional transport.

To make the multigrid scheme analogous to the angular multigrid method of [3], we have only used a semicoarsening strategy with coarsening in space only. We have not analyzed any multigrid scheme involving coarsening in time. However, analyzing a different coarsening strategy including coarsening in time is a natural extension of our LFA. This could be of interest in the case of extending the angular multigrid method to coarsen in angle and space simultaneously.

REFERENCES

[1] E. Abbott, *Flatland: A Romance of Many Dimensions*, Seely & Co., London, 1884.

[2] C. Börgers and E. Larsen, *On the accuracy of the Fokker-Planck and Fermi pencil beam equations for charged particle transport*, Medical Phys., 23 (1996), pp. 1749–59.

[3] C. Börgers and S. MacLachlan, *An angular multigrid method for computing mono-energetic particle beams in Flatland*, J. Comput. Phys., 229 (2010), pp. 2914–2931.

[4] C. Börgers, *A fast iterative method for computing particle beams penetrating matter*, J. Comput. Phys., 133 (1997), pp. 323–339.

[5] C. Börgers, *Complexity of Monte Carlo and deterministic dose-calculation methods*, Phys. Medicine Biology, 43 (1998), pp. 517–528.

[6] A. Brandt, *Multi-level adaptive solutions to boundary-value problems*, Math. Comp., 31 (1977), pp. 333–390.

[7] A. Brandt, *Guide to multigrid development*, in Multigrid Methods, Lecture Notes in Math. 960, Springer, Berlin, 1982, pp. 220–312.

[8] A. Brandt, *Rigorous quantitative analysis of multigrid.* I. *Constant coefficients two-level cycle with L_2-norm*, SIAM J. Numer. Anal., 31 (1994), pp. 1695–1730.

[9] R. Chan and M. Ng, *Conjugate gradient methods for Toeplitz systems*, SIAM Rev., 38 (1996), pp. 427–482.

[10] S. Friedhoff and S. MacLachlan, *A Predictive Analysis Tool for Multigrid Methods for Parabolic Problems*, Technical report, Tufts University, Medford, MA, 2013.

[11] L. Henyey and J. Greenstein, *Diffuse radiation in the galaxy*, Astrophys. J., 93 (1941), pp. 70–83.

[12] K. Hogstrom and P. Almond, *Review of electron beam therapy physics*, Phys. Medicine Biology, 51 (2006), pp. R455–R489.

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

[14] T. Ku and C.-C. J. Kuo, *Spectral properties of preconditioned rational Toeplitz matrices: The nonsymmetric case*, SIAM J. Matrix Anal. Appl., 14 (1993), pp. 512–544.

[15] E. Larsen and J. Morel, *Advances in discrete-ordinates methodology*, in Nuclear Computational Science: A Century in Review, Y. Y. Azmy and E. Sartori, eds., Springer-Verlag, Berlin, 2010.

[16] E. Lewis and W. Miller, *Computational Methods of Neutron Transport*, American Nuclear Society, La Grange Park, IL, 1993.
[17] C. Lubich and A. Ostermann, *Multi-grid dynamic iteration for parabolic equations*, BIT, 27 (1987), pp. 216–234.
[18] G. Pomraning, *The Fokker-Planck operator as an asymptotic limit*, Math. Models Methods Appl. Sci., 2 (1992), pp. 21–36.
[19] L. Reimer, *Scanning Electron Microscopy*, Springer-Verlag, Berlin, 1985.
[20] D. Rogers and A. Bielajew, *Monte Carlo techniques of electron and photon transport for radiation dosimetry*, in The Dosimetry of Ionizing Radiation, K. R. Kase, B. Bjärngard, and F. Attix, eds., Vol. 3, Academic Press, New York, 1990.
[21] R. Stevenson, *On the validity of local mode analysis of multi-grid methods*, Ph.D. thesis, University of Utrecht, Utrecht, The Netherlands, 1990.
[22] K. Stüben and U. Trottenberg, *Multigrid methods: Fundamental algorithms, model problem analysis and applications*, GMD-Studien 96, Gesellschaft für Mathematik und Datenverarbeitung mbH, St. Augustin, Germany, 1985.
[23] U. Trottenberg, C. Oosterlee, and A. Schüller, *Multigrid*, Academic Press, New York, 2001.
[24] S. Vandewalle and G. Horton, *Fourier mode analysis of the multigrid waveform relaxation and time-parallel multigrid methods*, Computing, 54 (1995), pp. 317–330.
[25] C. Zerby and F. Keller, *Electron transport theory, calculations, and experiments*, Nuclear Science and Engineering, 27 (1967), pp. 190–218.