SIAM REVIEW
Vol. 47, No. 3, pp. 447–468

A Fast and Stable Solution Method for the Radiative Transfer Problem*

Per Edström†

Abstract. Radiative transfer theory considers radiation in turbid media and is used in a wide range of applications. This paper outlines a problem formulation and a solution method for the radiative transfer problem in multilayer scattering and absorbing media using discrete ordinate model geometry. A selection of different steps is brought together. The main contribution here is the synthesis of these steps, all of which have been used in different areas, but never all together in one method. First, all necessary steps to get a numerically stable solution procedure are treated, and then methods are introduced to increase the speed by a factor of several thousand. This includes methods for handling strongly forward-scattering media. The method is shown to be unconditionally stable, though the problem was previously considered numerically intractable.

Key words. radiative transfer, discrete ordinates, solution method, numerical stability, speed

AMS subject classifications. 65R20, 85A25, 45J05

DOI. 10.1137/S0036144503438718

1. Introduction.

Radiative transfer theory describes the interaction of radiation with scattering and absorbing media. Radiative transfer is applied to such different areas of application as diffusion of neutrons, stellar atmospheres, optical tomography, infrared and visible light in space and the atmosphere, and light scattering from pigment films, paper, and print. Models for calculating the light intensity within and outside an illuminated turbid medium involve several numerical challenges and are crucial for a number of sectors of industry. Solution methods for radiative transfer problems have been studied throughout the last century.

In the beginning most radiative transfer problems were considered intractable because of numerical difficulties, so coarse approximations were used, and methods developed slowly due to the lack of mathematical tools. As computers have become faster and more readily available, highly efficient and specialized solution methods have been developed. Among the solution methods in use today are discrete ordinate methods (approximating integrals with numerical quadrature), methods using spherical harmonics (orthogonal functions), methods using finite elements or finite differences, and Monte-Carlo methods. This paper focuses on discrete ordinate methods only.

*Received by the editors December 17, 2003; accepted for publication (in revised form) July 22, 2004; published electronically July 29, 2005. This work was financially supported by the Swedish printing research program T2F, "TryckTeknisk Forskning."

http://www.siam.org/journals/sirev/47-3/43871.html

†Department of Engineering, Physics and Mathematics, Mid Sweden University, Gånsviksvägen 2, 87188 Härnösand, Sweden (per.edstrom@miun.se).

The first approximate solution was presented by Schuster [1], who considered only diffuse radiation, and exclusively in a forward and a backward direction. Clearly influenced by this, Kubelka and Munk [2] developed their model, well known in some applications, which was further refined by Kubelka [3, 4]. Despite several limitations, the Kubelka–Munk model is in widespread use for multiple scattering calculations in paper, coatings, printed paper, paint, plastic, and textile, probably due to its explicit form and ease of use. The models presented by Schuster and Kubelka and Munk, and others after them, are known as two-flux models.

By using numerical quadrature to approximate an integral with a finite sum, Wick [5] gave the first general treatment of discrete ordinate methods. The terms in the sum can be interpreted as the contribution to radiation from a discrete cone in spherical geometry. The polar angles of these cones are referred to as discrete ordinates, which has given the method its name, and the cones are called channels or streams. Using only two channels gives the earlier two-flux methods. If more channels are used, the methods are referred to as multiflux methods or many-flux methods.

Chandrasekhar described a method using spherical harmonics [6], but having read Wick's article, he adopted the discrete ordinate method and further refined it [7]. Later, he wrote a classic exposition on radiative transfer theory in book form [8], and since then the area has expanded tremendously.

Mudgett and Richards [9, 10] described a discrete ordinate method for use in technology and reported on numerical difficulties, as have many before and after them. These difficulties worsened when the use of computers made it possible to tackle larger problems. Only when recognizing the numerical difficulties can measures be taken. A careful analysis of the problem makes it possible to find such measures, and advances in numerical linear algebra and scientific computing provide ideas and software tools to make it a tractable problem. The point of this paper is not to describe the best possible solution method for the radiative transfer problem; the field is so diverse that specialized routines are needed that exploit the special properties of each specific application area. Instead, the point is to present a synthesis of the steps that are needed or possible to make *any* discrete ordinate radiative transfer solution method numerically efficient. To the author's knowledge, this has not been summarized in one single publication before.

2. Problem Formulation. For an ideally reflecting medium, all incoming light is specularly reflected at the surface. For a turbid medium, transmission as well as absorption and multiple scattering inside the medium have to be taken into consideration. In this paper, the problem is studied in a plane-parallel geometry, where the horizontal extension of the medium is assumed to be large enough to give no boundary effects at the sides. The boundary conditions at the top and bottom boundary surfaces, including illumination, are assumed to be time- and space-independent on the respective boundary surface. The radiation is assumed to be monochromatic, or confined to a narrow enough wavelength range to make scattering and absorption constant. The scattering is assumed to be conservative, i.e., without change in frequency between incoming and outgoing radiation. The medium is treated as a continuum of scattering and absorption sites. Polarization effects are ignored, hence using only the first component of the Stokes 4-vector. What is left is then a scalar intensity, which is the variable to solve for.

2.1. Some Definitions. The energy flow is thought of as noninteracting beams of radiation in all directions. This makes it possible to treat the beams separately. The intensity, I, of the radiation is always considered to be positive. When radiation

traverses a finite thickness ds of the medium in its direction of propagation, a fraction is extinct due to absorption and scattering. The intensity then becomes $I + dI$, and the extinction coefficient is defined as

$$\sigma_e = -\frac{dI}{Ids}.$$

The extinction coefficient can be separated into two parts, called the absorption and scattering coefficients, σ_a and σ_s, corresponding to the two different origins of the extinction. They are related to the extinction coefficient through $\sigma_e = \sigma_a + \sigma_s$. A convenient measure is the single scattering albedo, which is the probability for scattering given an extinction event, and is defined as

$$a = \frac{\sigma_s}{\sigma_e} = \frac{\sigma_s}{\sigma_a + \sigma_s}.$$

The phase function, p, specifies the angular distribution of the scattered radiation. If the phase function is normalized by

$$\int_0^{2\pi} \int_0^{\pi} \sin\theta \frac{p(\theta', \varphi'; \theta, \varphi)}{4\pi} d\theta d\varphi = 1,$$

where θ and φ are the polar and azimuthal angle coordinates of spherical geometry for the direction of the radiation (in the remainder of this paper, primed arguments correspond to incident radiation), this can be given a probabilistic interpretation. Given that radiation in the direction (θ', φ') is scattered, the probability that it is scattered into the cone of solid angle $d\theta d\varphi$ centered on the direction (θ, φ) is

$$\frac{p(\theta', \varphi'; \theta, \varphi) d\theta d\varphi}{4\pi}.$$

Different phase functions have been proposed to physically describe different types of scattering. Among the best known are the phase functions given by Rayleigh [11] and Mie [12]. The only one considered in this paper is the Henyey–Greenstein [13] phase function. It should not be seen as a real phase function, but is a one-parameter analytical approximation. It is given by

$$p(\cos\Theta) = \frac{1 - g^2}{(1 + g^2 - 2g\cos\Theta)^{3/2}}. \tag{2.1}$$

Here, Θ is the scattering angle, and it is evident that the Henyey–Greenstein phase function is dependent on the scattering angle Θ only, and not on the specific directions of incident and scattered radiation. The angular variables are related through the cosine law of spherical geometry as

$$\cos\Theta = \cos\theta' \cos\theta + \sin\theta' \sin\theta \cos(\varphi' - \varphi).$$

The coefficients for Legendre polynomial expansion (sometimes referred to as moments) of the Henyey–Greenstein phase function are simply $\chi_l = g^l$. The parameter g, here called the asymmetry factor, controls the scattering pattern, ranging from complete forward scattering ($g = 1$) over isotropic scattering ($g = 0$) to complete backward scattering ($g = -1$).

2.2. The Equation of Radiative Transfer. For a plane-parallel geometry, it is convenient to measure distances normal to the surface of the medium. This coincides with the z-axis in a Cartesian coordinate system if the surface is placed in the x-y-plane, and it is evident that $dz = ds\cos\theta$. The optical depth, measured from the top surface and down, is then defined as

$$\tau(z) = \int_z^\infty \sigma_e dz'.$$

It is also common to introduce $u = \cos\theta$, which gives $d\tau = -\sigma_e u ds$. Chandrasekhar [8, eq. I.71] states the equation of radiative transfer for a scattering plane-parallel medium as

$$u\frac{dI(\tau,u,\varphi)}{d\tau} = I(\tau,u,\varphi) - \frac{a}{4\pi}\int_0^{2\pi}\int_{-1}^{1} p(u',\varphi';u,\varphi)I(\tau,u',\varphi')du'd\varphi'. \tag{2.2}$$

The integral term on the right-hand side is a source function. It gives the intensity scattered from all incoming directions at a point to a specified direction. It is possible to add a term for emission, e.g., fluorescence or thermal emission, to the source function if the emission is inside the wavelength range of interest. These terms are easy to fit into the solution procedure. To perform the coupling of the intensities of the different wavelengths associated with the fluorescence, an outer loop over wavelengths will be needed. It should be noted here that the equations are not necessarily energy conservative, since absorbed light is always emitted as fluorescence or thermal radiation in wavelengths that may be ignored.

3. Solution Method. The intensity is described by an integrodifferential equation, the solution of which is the goal of this paper. The outline of the solution method is as follows. Fourier analysis gives a system of equations, which are then discretized using numerical quadrature. The initial problem can then be transferred to a problem on eigenvalues of matrices. Boundary and continuity conditions are imposed, and the computed intensity is extended from the quadrature points to the entire interval through interpolation formulas.

The main steps to achieve a numerically stable solution procedure include the Fourier analysis, the evaluation of normalized associated Legendre functions, the choice of numerical quadrature, the matrix formulation of the discretization, the reduction of the eigenvalue problem, the preconditioning of the system of equations corresponding to the boundary and continuity conditions, and the avoidance of over- and underflow in the solution and interpolation formulas. The recognition of potential divide-by-zero situations and reformulation of those are also important.

To make the method fast several measures are taken. The δ-N method and the intensity correction procedures allow high speed by maintaining accuracy at a significantly lower number of terms in the quadrature formula than would otherwise be needed. Computational shortcuts stop the calculations earlier when certain convergence criteria have been met. In addition, the sparse structure of the system of equations corresponding to the boundary and continuity conditions should be exploited.

3.1. Fourier Analysis on φ. The unknown intensity is a function of three variables, τ, u, and φ. It is possible to reduce the problem by factoring out the φ-dependence. This is achieved by Legendre function expansion of the phase function and then Fourier analysis on the azimuthal angle variable φ. This gives a set of radiative transfer equations that depend only on τ and u.

The key is to expand the phase function in a series of $2N$ Legendre polynomials as

$$p(\cos\Theta) \approx \sum_{l=0}^{2N-1} (2l+1)\chi_l P_l(\cos\Theta), \tag{3.1}$$

where $P_l(\cos\Theta)$ is the Legendre polynomial of degree l, and χ_l is the corresponding expansion coefficient. The Legendre polynomials are chosen for several reasons. They are a natural basis set of orthogonal polynomials on $[-1,1]$. Furthermore, they are used in Gaussian quadrature schemes for evaluating integrals numerically, and they give a simple expansion for the Henyey–Greenstein phase function. Finally, they enable separation of the angular coordinates u and φ through the addition theorem for spherical harmonics.

The addition theorem states that

$$P_l(\cos\Theta) = P_l(u')P_l(u) + 2\sum_{m=1}^{l} \Lambda_l^m(u')\Lambda_l^m(u)\cos(m(\varphi'-\varphi)),$$

where

$$\Lambda_l^m(u) = \sqrt{\frac{(l-m)!}{(l+m)!}} P_l^m(u)$$

are normalized associated Legendre functions and $P_l^m(u)$ are associated Legendre functions. The normalized functions are preferred since they remain bounded, while the nonnormalized functions can become large enough to cause overflow. The addition theorem allows the phase function, through the Legendre polynomial expansion, to be expressed as products of functions of u and φ separately. Introducing the function

$$p^m(u',u) = \sum_{l=m}^{2N-1} (2l+1)\chi_l \Lambda_l^m(u')\Lambda_l^m(u),$$

the phase function can be expressed as

$$p(u',\varphi';u,\varphi) = \sum_{m=0}^{2N-1} (2-\delta_{0m}) p^m(u',u)\cos(m(\varphi'-\varphi)). \tag{3.2}$$

This is in essence a Fourier cosine series for the phase function, and it makes sense to expand the intensity in a similar way:

$$I(\tau,u,\varphi) = \sum_{m=0}^{2N-1} I^m(\tau,u)\cos(m(\varphi_0-\varphi)), \tag{3.3}$$

where I^m are the Fourier components of the intensity and φ_0 is some suitably chosen reference. Inserting these Fourier cosine series expansions for the phase function and for the intensity into the equation of radiative transfer (2.2), the integral term after some rearrangements becomes

$$\begin{aligned}&\frac{a}{4\pi}\int_0^{2\pi}\int_{-1}^{1} p(u',\varphi';u,\varphi) I(\tau,u',\varphi')\,du'\,d\varphi' \\ &= \frac{a}{2}\sum_{m=0}^{2N-1}\cos(m(\varphi_0-\varphi))\int_{-1}^{1} p^m(u',u) I^m(\tau,u')\,du'.\end{aligned}$$

This gives an equation for each of the Fourier components as

$$u\frac{dI^m(\tau,u)}{d\tau} = I^m(\tau,u) - \frac{a}{2}\int_{-1}^{1} p^m(u',u)I^m(\tau,u')du', \qquad m = 0,\ldots,2N-1. \tag{3.4}$$

These equations are entirely uncoupled and can be solved independently. The complete azimuthal dependence can then be assembled through the Fourier cosine series expansion for the intensity above. Thus, the dependence of the variable φ is totally eliminated.

3.2. Enhancing Symmetry. Half-range intensities are now introduced to exploit the symmetry of the problem. They are denoted I^+ and I^-, where the plus and minus signs designate intensities in the upper and lower hemispheres, i.e., for $0 \le \theta \le \pi/2$ and $\pi/2 < \theta \le \pi$, respectively. It is also beneficial to use $\mu = |u| = |\cos\theta|$. Furthermore, most relevant illumination conditions are either diffuse, a directed beam, or a combination of both. Therefore it is convenient to separate the intensity into the corresponding components, a diffuse component I_d, and a beam component I_b. The beam component is assumed to be infinitesimally narrow, so it suffers from absorption but it does not get any contribution from scattering from other directions. Therefore, it is simply

$$I_b^-(\tau,\mu,\varphi) = I_{0b}e^{-\tau/\mu_0}\delta(\mu-\mu_0)\delta(\varphi-\varphi_0), \tag{3.5}$$

where I_{0b} and (μ_0,φ_0) are the intensity and direction of the incident beam and δ is the Dirac delta function. The diffuse component is also called the multiple-scattering component and includes reflection from the bottom boundary surface. The beam component is therefore present in downward directions only, $I^- = I_d^- + I_b^-$, but not in upward directions, $I^+ = I_d^+$. Using these expressions for I^+ and I^- (now dropping the subscript d) yields the following pair of coupled integrodifferential equations for the Fourier components of the diffuse intensity, since the nonintegral terms involving I_b^- cancel:

$$\begin{cases} \mu\dfrac{dI^{m+}(\tau,\mu)}{d\tau} = I^{m+}(\tau,\mu) - \dfrac{a}{2}\displaystyle\int_0^1 p^m(\mu',\mu)I^{m+}(\tau,\mu')d\mu' \\ \qquad\qquad - \dfrac{a}{2}\displaystyle\int_0^1 p^m(-\mu',\mu)I^{m-}(\tau,\mu')d\mu' - X_0^{m+}e^{-\tau/\mu_0}, \\ -\mu\dfrac{dI^{m-}(\tau,\mu)}{d\tau} = I^{m-}(\tau,\mu) - \dfrac{a}{2}\displaystyle\int_0^1 p^m(\mu',-\mu)I^{m+}(\tau,\mu')d\mu' \\ \qquad\qquad - \dfrac{a}{2}\displaystyle\int_0^1 p^m(-\mu',-\mu)I^{m-}(\tau,\mu')d\mu' - X_0^{m-}e^{-\tau/\mu_0}, \end{cases} \qquad m = 0,\ldots,2N-1, \tag{3.6}$$

where

$$X_0^{m\pm} = \frac{a}{4\pi}(2-\delta_{0m})p^m(-\mu_0,\pm\mu)I_{0b}. \tag{3.7}$$

3.3. Evaluation of the Normalized Associated Legendre Functions. There are many ways of evaluating associated Legendre functions numerically, and a lot of them are poor. For example, explicit expressions involve cancellation between successive

terms, which alternate in sign. For large l, the individual terms become larger than their sum, and all accuracy is lost.

The associated Legendre functions satisfy a number of recurrence relations on either or both of l and m. Most of the recurrences on m are unstable and hence numerically unsuitable. This paper uses the following three-term recurrence on l from Magnus and Oberhettinger [14], which is stable:

$$(l-m)P_l^m(u) = u(2l-1)P_{l-1}^m(u) - (l-1+m)P_{l-2}^m(u).$$

From this, the recurrence for the normalized functions can be found to be

$$\Lambda_l^m(u) = \frac{u(2l-1)\Lambda_{l-1}^m(u) - \sqrt{(l-1+m)(l-1-m)}\Lambda_{l-2}^m(u)}{\sqrt{(l-m)(l+m)}}. \tag{3.8}$$

The three-term recurrence for the associated Legendre functions has a closed-form expression for the starting value,

$$P_m^m(u) = (-1)^m(2m-1)!!(1-u^2)^{m/2}.$$

This can be translated into a two-term recurrence for the normalized functions,

$$\begin{cases} \Lambda_0^0(u) = 1, \\ \Lambda_m^m(u) = -\sqrt{1-u^2}\sqrt{\frac{2m-1}{2m}}\Lambda_{m-1}^{m-1}(u). \end{cases} \tag{3.9}$$

If the three-term recurrence for the associated Legendre functions is used with $l = m+1$, and using the convention $P_{m-1}^m(u) = 0$, the result is

$$P_{m+1}^m(u) = u(2m+1)P_m^m(u).$$

For the normalized functions, this becomes

$$\Lambda_{m+1}^m(u) = u\sqrt{2m+1}\Lambda_m^m(u). \tag{3.10}$$

All together, this constitutes a numerically stable way to compute the normalized associated Legendre functions.

3.4. Double-Gauss Quadrature. One problem in radiative transfer is to calculate integrals of the form

$$\int_{-1}^{1} f(u)du.$$

This integral can be approximated by a finite sum, a numerical quadrature formula, as

$$\int_{-1}^{1} f(u)du \approx \sum_{j=1}^{m} \omega_j' f(u_j).$$

Different choices of the weights ω_j' and nodes u_j give different quadrature formulas. If the nodes are taken linearly spaced from -1 to 1, there is a unique choice of weights that gives the quadrature an order of accuracy of at least $m-1$. This is known

as a Newton–Cotes formula. It is simple and useful for small m, but for larger m, their weights have oscillating signs and amplitudes of the order of 2^m, which causes numerical instability. Gauss showed that if not only the weights but also the nodes are chosen optimally, the result is a formula of order $2m-1$, which is the best possible. This is known as a Gaussian quadrature formula. The optimal nodes are the zeros of the Legendre polynomial $P_m(u)$. Furthermore, the weights are all positive, which makes the formula numerically stable even for large m.

There are closed expressions for the coefficients in the Legendre polynomials, but there is a risk of overflow for large m. The Lanczos iteration is a numerically stable method for finding the Legendre polynomial coefficients, but it is still unstable to find zeros directly from polynomial coefficients. However, there is a closed expression for the Jacobi matrix used in the Lanczos iteration. By solving an eigenvalue problem for the Jacobi matrix, the optimal weights and nodes can be found without even forming the Legendre polynomials, as suggested by Golub and Welsch [15]. The eigenvalues of the Jacobi matrix are the required nodes, and the weights are twice the square of the first component of the eigenvectors. Thus, this is a fast and stable method for finding the nodes and weights for a quadrature formula with optimal accuracy. Furthermore, there is an advantage in using Gaussian quadrature of even order; symmetry ensures that the nodes occur in pairs and that the corresponding weights are equal.

Gaussian quadrature assumes that the integrand is a smooth function. It is known, however, that the intensity changes rapidly close to $u = 0$ near the boundaries. Furthermore, Gaussian quadrature has the nodes the least dense close to $u = 0$, where the intensity changes the most. In order to improve the situation, a modification to the Gaussian quadrature is used.

Double-Gauss, proposed by Sykes [16], approximates the integral over the two hemispheres separately,

$$\int_{-1}^{1} f(u)du = \int_{0}^{1} f^{+}(\mu)d\mu + \int_{0}^{1} f^{-}(\mu)d\mu \approx \sum_{j=1}^{N} \omega_j f^{+}(\mu_j) + \sum_{j=1}^{N} \omega_j f^{-}(\mu_j),$$

where the nodes μ_j and weights ω_j are chosen for the "half interval" $[0, 1]$. For the greatest accuracy, the optimal Gaussian quadrature should be used on the new interval $0 \leq \mu \leq 1$, so with a simple translation, the Jacobi matrix from the Lanczos iteration can still be used to find μ_j and ω_j.

It should be noted that the N used here is the same one that was introduced for the phase function expansion in section 3.1. The correspondence of an expansion in $2N$ Legendre polynomials (which are the eigenfunctions of the scattering operator for the $m = 0$ equation) and a $2N$ point double-Gauss quadrature is important, since fewer points would not give photon conservation. Actually, to maintain optimal accuracy for Fourier components $m > 0$, different quadrature sets for each m that are specifically designed to integrate the associated Legendre functions (which are the eigenfunctions of the scattering operator for the $m > 0$ equations) would be needed. However, this would complicate the solution procedure a great deal and is thus not done here.

3.5. Matrix Formulation. The discrete ordinate approximation, i.e., application of the double-Gauss quadrature rule described above, can now be used to transform these pairs of coupled integrodifferential equations into systems of coupled ordinary differential equations. For each Fourier component (where the superscript m has been

dropped), this yields

$$
\left\{
\begin{aligned}
\mu_i \frac{dI^+(\tau,\mu_i)}{d\tau} &= I^+(\tau,\mu_i) - \frac{a}{2}\sum_{j=1}^{N}\omega_j p(\mu_j,\mu_i) I^+(\tau,\mu_j) \\
&\quad - \frac{a}{2}\sum_{j=1}^{N}\omega_j p(-\mu_j,\mu_i) I^-(\tau,\mu_j) - X_{0i}^+ e^{-\tau/\mu_0}, \\
-\mu_i \frac{dI^-(\tau,\mu_i)}{d\tau} &= I^-(\tau,\mu_i) - \frac{a}{2}\sum_{j=1}^{N}\omega_j p(\mu_j,-\mu_i) I^+(\tau,\mu_j) \\
&\quad - \frac{a}{2}\sum_{j=1}^{N}\omega_j p(-\mu_j,-\mu_i) I^-(\tau,\mu_j) - X_{0i}^- e^{-\tau/\mu_0}, \\
& \qquad i = 1,\ldots,N.
\end{aligned}
\right. \tag{3.11}
$$

This was suggested by Stamnes and Swanson [17], who also put it in matrix form as

$$
\frac{d}{d\tau}\begin{bmatrix} \mathbf{I}^+ \\ \mathbf{I}^- \end{bmatrix} = \begin{bmatrix} -\alpha & -\beta \\ \beta & \alpha \end{bmatrix}\begin{bmatrix} \mathbf{I}^+ \\ \mathbf{I}^- \end{bmatrix} - \begin{bmatrix} \mathbf{Q}^+ \\ \mathbf{Q}^- \end{bmatrix}, \tag{3.12}
$$

where

$$
\begin{aligned}
\mathbf{I}^{\pm} &= \{I^{\pm}(\tau,\mu_i)\}, && i = 1,\ldots,N, \\
\mathbf{Q}^{\pm} &= \pm\mathbf{M}^{-1}\mathbf{Q}'^{\pm} = \{Q^{\pm}(\tau,\mu_i)\}, && i = 1,\ldots,N, \\
\mathbf{Q}'^{\pm} &= \left\{\frac{a}{4\pi}(2-\delta_{0m})p^m(-\mu_0,\pm\mu_i)I_{0b}e^{-\tau/\mu_0}\right\}, && i = 1,\ldots,N, \\
\mathbf{M} &= \{\mu_i\delta_{ij})\}, && i,j = 1,\ldots,N, \\
\alpha &= \mathbf{M}^{-1}\left(\frac{a}{2}\mathbf{D}^+\mathbf{W} - \mathbf{1}\right), \\
\beta &= \mathbf{M}^{-1}\frac{a}{2}\mathbf{D}^-\mathbf{W}, \\
\mathbf{W} &= \{\omega_i\delta_{ij})\}, && i,j = 1,\ldots,N, \\
\mathbf{1} &= \{\delta_{ij})\}, && i,j = 1,\ldots,N, \\
\mathbf{D}^+ &= \{p(\mu_j,\mu_i)\} = \{p(-\mu_j,-\mu_i)\}, && i,j = 1,\ldots,N, \\
\mathbf{D}^- &= \{p(-\mu_j,\mu_i)\} = \{p(\mu_j,-\mu_i)\}, && i,j = 1,\ldots,N,
\end{aligned}
$$

and δ_{ij} is the Kronecker delta. It should be noted that this matrix formulation is identical for all Fourier components $m = 1,\ldots,2N-1$ if one simply replaces $p(\mu',\mu)$ for each m with $p^m(\mu',\mu)$.

3.6. Eigenvalue Problem. It is well known that the homogeneous solutions to systems of coupled ordinary differential equations such as (3.12) are of the form $\mathbf{I}^{\pm} = \mathbf{g}^{\pm}e^{-k\tau}$. This gives the eigenvalue problem

$$
\begin{bmatrix} \alpha & \beta \\ -\beta & -\alpha \end{bmatrix}\begin{bmatrix} \mathbf{g}^+ \\ \mathbf{g}^- \end{bmatrix} = k\begin{bmatrix} \mathbf{g}^+ \\ \mathbf{g}^- \end{bmatrix} \tag{3.13}
$$

of the size $2N\times 2N$ for the eigenvalues k and the eigenvectors $\mathbf{g}^{\pm}$. The structure of the $2N\times 2N$ matrix is due to the choice of numerical quadrature where the nodes come in pairs and the corresponding weights are equal, but it is also due to the phase function

being dependent on the scattering angle Θ only (so that the φ-dependence could be factored out). This structure ensures that the eigenvalues occur in positive/negative pairs, which allows reduction in the size of the eigenvalue problem by a factor of 2, and thus reduction in the eigenvalue calculations roughly by a factor of 8. This was noted already by Chandrasekhar [8], and Stamnes and Swanson [17] proposed the following solution to the eigenvalue problem. Adding and subtracting lines in (3.12) without $\mathbf{Q}^{\pm}$ and inserting the proposed homogeneous solutions $\mathbf{I}^{\pm} = \mathbf{g}^{\pm}e^{-k\tau}$ gives

$$(\alpha - \beta)(\alpha + \beta)(\mathbf{g}^{+} + \mathbf{g}^{-}) = k^2(\mathbf{g}^{+} + \mathbf{g}^{-}).$$

This is an eigenvalue problem for the eigenvectors $(\mathbf{g}^{+} + \mathbf{g}^{-})$ and the eigenvalues k^2 of size $N \times N$, i.e., half the original size. Finding $(\mathbf{g}^{+} - \mathbf{g}^{-})$ with some algebraic rearrangements and taking the sum and difference of $(\mathbf{g}^{+} + \mathbf{g}^{-})$ and $(\mathbf{g}^{+} - \mathbf{g}^{-})$ then gives the eigenvectors $\mathbf{g}^{\pm}$ for the original homogeneous eigenvalue problem.

It can be verified by insertion that

$$I(\tau, u_i) = Z_0(u_i)e^{-\tau/\mu_0} \tag{3.14}$$

is a particular solution if $Z_0(u_i)$ is determined by the system of linear equations

$$\sum_{\substack{-N \le j \le N \\ j \neq 0}} \left(\left(1 + \frac{u_j}{\mu_0} \right) \delta_{ij} - \omega_j \frac{a}{2} p(u_j, u_i) \right) Z_0(u_j) = X_0(u_i). \tag{3.15}$$

The general solution is given by the sum of the particular solution and a linear combination of the eigensolutions as

$$I^{\pm}(\tau, \mu_i) = \sum_{j=1}^{N} C_{-j} g_{-j}(\pm\mu_i) e^{k_j \tau} + \sum_{j=1}^{N} C_j g_j(\pm\mu_i) e^{-k_j \tau} + Z_0(\pm\mu_i) e^{-\tau/\mu_0}, \quad i = 1, \ldots, N. \tag{3.16}$$

Here, $\pm k_j$ and $g_{\pm j}(\pm\mu_i)$ are eigenvalues and eigenvectors, $\pm\mu_i$ are quadrature points, and $C_{\pm j}$ are constants to be given by boundary conditions. Also, $k_j > 0$ for positive j, and $k_{-j} = -k_j$.

These solutions pertain to a single, vertically homogeneous layer. There are at least two reasons for considering multilayer structures. One is that the medium might in fact be constructed as several discrete and vertically homogeneous layers placed on top of each other. Another is that an inhomogeneous medium can be approximated with a (sufficiently large) number of adjacent homogeneous layers. A method to handle refraction and total reflection at the boundaries of layers with different indices of refraction has been described by Jin and Stamnes [18] and can be included in this solution method if desired.

Since each of the layers in the multilayer structure is homogeneous—whether it is a real discrete structure or an approximation of a continuously varying one—the previously derived single layer solution can be used. Thus, the solution for the pth layer can be written as

$$I_p^{\pm}(\tau, \mu_i) = \sum_{j=1}^{N} \left(C_{jp} g_{jp}(\pm\mu_i) e^{-k_{jp}\tau} + C_{-jp} g_{-jp}(\pm\mu_i) e^{+k_{jp}\tau} \right) + U_p^{\pm}(\tau, \mu_i), \quad p = 1, \ldots, L, \tag{3.17}$$

where the sum is the homogeneous solution, $U_p^{\pm}(\tau, \mu_i)$ is the particular solution, and L is the number of layers. The only difference from the single layer case is the addition of the layer index p.

It should be noted that this is the solution for one Fourier component of the diffuse intensity. The complete azimuthal dependence can be assembled through the Fourier cosine series expansion for the diffuse intensity, as stated earlier. As a special case, the $m = 0$ component alone gives the azimuthal average, which is something several standardized measurements give, e.g., diffuse reflectance measurements.

It should also be noted that this eigenvalue problem can be formulated and solved in an alternative way, known in the neutron transport community as the method of separation of variables, which has been extensively developed by Barros and coworkers [19, 20, 21]. The interested reader should also be aware of the review by Badruzzaman [22], which discusses a number of relevant methods used in neutron transport problems.

3.7. Boundary and Continuity Conditions with Preconditioning. The interaction of the radiation with the bottom boundary surface of the medium (or with the surface of an underlying medium) can be described by a function, $\rho(-\mu', \varphi'; \mu, \varphi)$, that works in a similar manner as the phase function. The upward diffuse intensity at the bottom boundary surface is obtained by integrating over all the incident downward directions:

$$I_L^+(\tau_L, \mu, \varphi) = \frac{1}{\pi} \int_0^{2\pi} \int_0^1 \mu' \rho(-\mu', \varphi'; \mu, \varphi) I_L^-(\tau_L, \mu', \varphi') d\mu' d\varphi' + \frac{\mu_0}{\pi} \rho(-\mu_0, \varphi_0; \mu, \varphi) I_{0b} e^{-\tau_L/\mu_0},$$

where τ_L is the optical depth at the bottom boundary.

If $\rho(-\mu', \varphi'; \mu, \varphi)$ is assumed to depend on the difference $\varphi' - \varphi$, and not on the specific azimuthal directions of incident and reflected radiation, it can be expanded in a Fourier cosine series, and the azimuthal dependence can be factored out. Thus,

$$\rho(-\mu', \varphi'; \mu, \varphi) = \rho(-\mu', \mu; \varphi' - \varphi) = \sum_{m=0}^{2N-1} \rho^m(-\mu', \mu) \cos(m(\varphi' - \varphi)), \tag{3.18}$$

where

$$\rho^m(-\mu', \mu) = \frac{1}{\pi} \int_{-\pi}^{\pi} \rho(-\mu', \mu; \varphi' - \varphi) \cos(m(\varphi' - \varphi)) d(\varphi' - \varphi).$$

Using the Fourier cosine series expansions for both the diffuse intensity and ρ yields

$$\frac{1}{\pi} \int_0^{2\pi} \int_0^1 \mu' \rho(-\mu', \varphi'; \mu, \varphi) I_L^-(\tau_L, \mu', \varphi') d\mu' d\varphi' = \sum_{m=0}^{2N-1} (1 + \delta_{0m}) \cos(m(\varphi_0 - \varphi)) \int_0^1 \mu' \rho^m(-\mu', \mu) I_L^{m-}(\tau_L, \mu') d\mu'.$$

This gives the following condition for each Fourier component at the bottom boundary:

$$I_L^{m+}(\tau_L, \mu) = (1 + \delta_{0m}) \int_0^1 \mu' \rho^m(-\mu', \mu) I_L^{m-}(\tau_L, \mu') d\mu' + \frac{\mu_0}{\pi} \rho^m(-\mu_0, \mu) I_{0b} e^{-\tau_L/\mu_0},$$
$$m = 0, \ldots, 2N - 1.$$

But the multilayer solution contains $2N \times L$ constants to be determined, so in addition to boundary conditions, the intensity must also be required to be continuous across layer interfaces. Stamnes and Conklin [23] gave the formulation below of the problem of finding the unknown constants C_{jp}.

The conditions can be stated as a system of equations as (without the superscript m)

$$\begin{cases} I_1(0,-\mu_i) = \mathcal{I}(-\mu_i),\ i = 1,\dots,N, \\ I_p(\tau_p,\mu_i) = I_{p+1}(\tau_p,\mu_i),\ i = \pm 1,\dots,\pm N,\ p = 1,\dots,L-1, \\ I_L(\tau_L,+\mu_i) = (1+\delta_{m0}) \displaystyle\sum_{j=1}^{N} \omega_j \mu_j \rho(-\mu_j,\mu_i) I(\tau_L,-\mu_j) \\ \qquad\qquad + \dfrac{\mu_0}{\pi} \rho(-\mu_0,\mu_i) I_{0b} e^{-\tau_L/\mu_0},\ i = 1,\dots,N, \end{cases} \tag{3.19}$$

where $\mathcal{I}(-\mu_i)$ is the incident intensity at the top boundary surface. Inserting the multilayer solution (3.17) into this system of equations gives

$$\begin{cases} \displaystyle\sum_{j=1}^{N} \left(C_{j1} g_{j1}(-\mu_i) + C_{-j1} g_{-j1}(-\mu_i) \right) = \mathcal{I}(-\mu_i) - U_1(0,-\mu_i), \\ \qquad i = 1,\dots,N, \\ \displaystyle\sum_{j=1}^{N} \left\{ \left(C_{jp} g_{jp}(\mu_i) e^{-k_{jp}\tau_p} + C_{-jp} g_{-jp}(\mu_i) e^{+k_{jp}\tau_p} \right) \right. \\ \qquad \left. - \left(C_{j,p+1} g_{j,p+1}(\mu_i) e^{-k_{j,p+1}\tau_p} + C_{-j,p+1} g_{-j,p+1}(\mu_i) e^{+k_{j,p+1}\tau_p} \right) \right\} \\ \qquad = U_{p+1}(\tau_p,\mu_i) - U_p(\tau_p,\mu_i), \\ \qquad i = \pm 1,\dots,\pm N,\ p = 1,\dots,L-1, \\ \displaystyle\sum_{j=1}^{N} \left(C_{jL} r_j(\mu_i) e^{-k_{jL}\tau_L} + C_{-jL} r_{-j}(\mu_i) e^{+k_{jL}\tau_L} \right) = \Gamma(\tau_L,\mu_i), \\ \qquad i = 1,\dots,N, \end{cases} \tag{3.20}$$

where

$$r_j(\mu_i) = g_{jL}(+\mu_i) - (1+\delta_{m0}) \sum_{n=1}^{N} \rho(-\mu_n,\mu_i) \omega_n \mu_n g_{jL}(-\mu_n)$$

and

$$\begin{aligned} \Gamma(\tau_L,\mu_i) = &- U_L^+(\tau_L,\mu_i) + (1+\delta_{m0}) \sum_{j=1}^{N} \rho(-\mu_j,\mu_i) \omega_j \mu_j U_L^-(\tau_L,\mu_j) \\ &+ \frac{\mu_0}{\pi} \rho(-\mu_0,\mu_i) I_{0b} e^{-\tau_L/\mu_0}. \end{aligned}$$

The boundary and continuity conditions give a $(2N \times L) \times (2N \times L)$ system of equations for the $2N \times L$ unknown coefficients $C_{jp},\ j = \pm 1,\dots,\pm N,\ p = 1,\dots,L$. The coefficient matrix is sparse and block diagonal, with $6N-1$ diagonals, a fact that should be exploited in a numerical implementation.

However, the equations are ill-conditioned due to the exponentials with positive arguments. This is why the method was discarded in the past. But the ill-conditioning can be removed by using as a preconditioner the scaling transformation

$$C_{+jp} = C'_{+jp} e^{k_{jp}\tau_{p-1}} \quad \text{and} \quad C_{-jp} = C'_{-jp} e^{-k_{jp}\tau_p}, \tag{3.21}$$

where τ_p is the optical depth at the bottom of layer p. The scaled system of equations for the coefficients C'_{jp} then becomes (with τ_0 as the optical depth at the top)

$$\begin{cases} \displaystyle\sum_{j=1}^{N} \left(C'_{j1} g_{j1}(-\mu_i) + C'_{-j1} g_{-j1}(-\mu_i) e^{-k_{jp}(\tau_1 - \tau_0)} \right) = \mathcal{I}(-\mu_i) - U_1(0, -\mu_i), \\ \quad i = 1, \ldots, N, \\ \displaystyle\sum_{j=1}^{N} \Big\{ \left(C'_{jp} g_{jp}(\mu_i) e^{-k_{jp}(\tau_p - \tau_{p-1})} + C'_{-jp} g_{-jp}(\mu_i) \right) \\ \quad - \left(C'_{j,p+1} g_{j,p+1}(\mu_i) + C'_{-j,p+1} g_{-j,p+1}(\mu_i) e^{-k_{j,p+1}(\tau_{p+1} - \tau_p)} \right) \Big\} \\ \quad = U_{p+1}(\tau_p, \mu_i) - U_p(\tau_p, \mu_i), \\ \quad i = \pm 1, \ldots, \pm N, \ p = 1, \ldots, L-1, \\ \displaystyle\sum_{j=1}^{N} \left(C'_{jL} r_j(\mu_i) e^{-k_{jL}(\tau_L - \tau_{L-1})} + C'_{-jL} r_{-j}(\mu_i) \right) = \Gamma(\tau_L, \mu_i), \\ \quad i = 1, \ldots, N. \end{cases} \tag{3.22}$$

Since $k_{jp} > 0$ and $\tau_p > \tau_{p-1}$, all exponentials in the system of equations for the coefficients C'_{jp} have negative arguments. Thus, the ill-conditioning is prevented, and the problem of solving for the C'_{jp} is unconditionally stable.

There is a risk of overflow when evaluating the solution for the pth layer, but this can be avoided with the use of the coefficients C'_{jp}. By using the same scaling as with the boundary and continuity conditions, the general solution becomes

$$I_p^{\pm}(\tau, \mu_i) = \sum_{j=1}^{N} (C'_{jp} g_{jp}(\pm\mu_i) e^{-k_{jp}(\tau - \tau_{p-1})} + C'_{-jp} g_{-jp}(\pm\mu_i) e^{-k_{jp}(\tau_p - \tau)}) + U_p^{\pm}(\tau, \mu_i), \tag{3.23}$$
$$p = 1, \ldots, L.$$

Since $k_{jp} > 0$ and $\tau_{p-1} < \tau < \tau_p$, all exponentials have negative arguments, and the risk of overflow is prevented. It should be pointed out that in a numerical implementation the scaled coefficients should be used in the rest of the solution procedure, which makes any rescaling transformation unnecessary and thus eliminates the risk of enlarging errors later.

3.8. Interpolation Formulas. The general solution for the discrete problem gives the intensity at any depth, but only in the quadrature points. If the intensity in an arbitrary direction is required, interpolation formulas are needed. It is always possible to fit a polynomial to a number of points. A polynomial of sufficiently high degree will be exact in all points to be fitted, but will normally perform badly between the points. If a polynomial of lower degree is chosen, it will perform better between the points to be fitted, but on the other hand it will not be exact in those points, even

though they are known. It is also possible to use cubical splines. They will be exact in all points to be fitted, but they will also perform badly between if there are large changes in one or more of the points. Another approach is to use the solutions of the eigenvalue problem, as proposed by Stamnes [24], and that scheme is outlined below.

Although derived for a single layer, the discrete equations for the Fourier components (3.11) are equally valid across all layers together. They can, substituting the quadrature points μ_i for the free variable μ, be written

$$\begin{cases} \mu \dfrac{dI^+(\tau,\mu)}{d\tau} = I^+(\tau,\mu) - S^+(\tau,\mu), \\ -\mu \dfrac{dI^-(\tau,\mu)}{d\tau} = I^-(\tau,\mu) - S^-(\tau,\mu), \end{cases} \tag{3.24}$$

where

$$\begin{cases} S^+(\tau,\mu) = \dfrac{a}{2}\displaystyle\sum_{i=1}^{N} \omega_i p(\mu_i,\mu) I^+(\tau,\mu_i) \\ \qquad\qquad + \dfrac{a}{2}\displaystyle\sum_{i=1}^{N} \omega_i p(-\mu_i,\mu) I^-(\tau,\mu_i) + X_0^+(\mu) e^{-\tau/\mu_0}, \\ S^-(\tau,\mu) = \dfrac{a}{2}\displaystyle\sum_{i=1}^{N} \omega_i p(\mu_i,-\mu) I^+(\tau,\mu_i) \\ \qquad\qquad + \dfrac{a}{2}\displaystyle\sum_{i=1}^{N} \omega_i p(-\mu_i,-\mu) I^-(\tau,\mu_i) + X_0^-(\mu) e^{-\tau/\mu_0}, \end{cases}$$

provided proper layer indexing—depending on the optical depth considered—is used throughout.

Inserting the multilayer solution (3.17) into this expression for the source functions yields

$$S_p^{\pm}(\tau,\mu) = \sum_{j=1}^{N} C_{-jp} \tilde{g}_{-jp}(\pm\mu) e^{k_{jp}\tau} + \sum_{j=1}^{N} C_{jp} \tilde{g}_{jp}(\pm\mu) e^{-k_{jp}\tau} + \widetilde{Z}_{0p}^{\pm}(\mu) e^{-\tau/\mu_0}, \tag{3.25}$$

where

$$\tilde{g}_{jp}(\pm\mu) = \frac{a}{2}\sum_{i=1}^{N} \big(\omega_i p(-\mu_i,\pm\mu) g_{jp}(-\mu_i) + \omega_i p(+\mu_i,\pm\mu) g_{jp}(+\mu_i)\big)$$

and

$$\widetilde{Z}_{0p}^{\pm}(\mu) = \frac{a}{2}\sum_{i=1}^{N} \big(\omega_i p(-\mu_i,\pm\mu) Z_{0p}(-\mu_i) + \omega_i p(+\mu_i,\pm\mu) Z_{0p}(+\mu_i)\big) + X_{0p}(\pm\mu).$$

These are analytical interpolation formulas for the source function for each layer, expressed in the solutions of the eigenvalue problem for each respective layer.

Equations (3.24) can be integrated formally, giving analytical formulas for the intensity at arbitrary depth and direction expressed in the source function. Inserting the interpolation formulas for the source function (3.25) then gives interpolation formulas for the intensity as well, thus making it possible to calculate the intensity at any depth and at any angle.

As with the discrete solution (3.17), there is a risk of overflow when evaluating the interpolation formulas, but this can be avoided by the use of the coefficients C'_{jp}. By using the same scaling as with the boundary and continuity conditions, the interpolation formulas become

(3.26)
$$\begin{aligned} I_p^+(\tau,\mu) = {} & I_L^+(\tau_L,\mu)e^{-(\tau_L-\tau)/\mu} \\ & + \sum_{n=p}^{L} \left\{ \frac{\widetilde{Z}_{0p}(+\mu)}{1+\mu/\mu_0} \left(e^{-(k_{jn}\tau_{n-1}+(\tau_{n-1}-\tau)/\mu)} - e^{-(k_{jn}\tau_n+(\tau_n-\tau)/\mu)} \right) \right. \\ & \quad + \sum_{j=1}^{N} C'_{jn} \frac{\tilde{g}_{jn}(+\mu)}{1+k_{jn}\mu} \left(e^{-(\tau_{n-1}-\tau)/\mu} - e^{-(k_{jn}(\tau_n-\tau_{n-1})+(\tau_n-\tau)/\mu)} \right) \\ & \quad \left. + \sum_{j=1}^{N} C'_{-jn} \frac{\tilde{g}_{-jn}(+\mu)}{1-k_{jn}\mu} \left(e^{-(k_{jn}(\tau_n-\tau_{n-1})+(\tau_{n-1}-\tau)/\mu)} - e^{-(\tau_n-\tau)/\mu} \right) \right\} \end{aligned}$$

with τ_{n-1} replaced by τ and the exponentials in the second sum replaced by

$$e^{-k_{jp}(\tau-\tau_{p-1})} - e^{-(k_{jp}(\tau_p-\tau_{p-1})+(\tau_p-\tau)/\mu)}$$

for $n = p$, and

(3.27)
$$\begin{aligned} I_p^-(\tau,\mu) = {} & I_0^-(\tau_0,\mu)e^{-(\tau-\tau_0)/\mu} \\ & + \sum_{n=1}^{p} \left\{ \frac{\widetilde{Z}_{0p}(-\mu)}{1-\mu/\mu_0} \left(e^{-(k_{jn}\tau_n+(\tau-\tau_n)/\mu)} - e^{-(k_{jn}\tau_{n-1}+(\tau-\tau_{n-1})/\mu)} \right) \right. \\ & \quad + \sum_{j=1}^{N} C'_{jn} \frac{\tilde{g}_{jn}(-\mu)}{1-k_{jn}\mu} \left(e^{-(k_{jn}(\tau_n-\tau_{n-1})+(\tau-\tau_n)/\mu)} - e^{-(\tau-\tau_{n-1})/\mu} \right) \\ & \quad \left. + \sum_{j=1}^{N} C'_{-jn} \frac{\tilde{g}_{-jn}(-\mu)}{1+k_{jn}\mu} \left(e^{-(\tau-\tau_n)/\mu} - e^{-(k_{jn}(\tau_n-\tau_{n-1})+(\tau-\tau_{n-1})/\mu)} \right) \right\} \end{aligned}$$

with τ_n replaced by τ and the exponentials in the third sum replaced by

$$e^{-k_{jp}(\tau_p-\tau)} - e^{-(k_{jp}(\tau_p-\tau_{p-1})+(\tau-\tau_{p-1})/\mu)}$$

for $n = p$. Since all $k_{jn} > 0$ (and especially $k_{jp} > 0$) and $\tau_{p-1} < \tau < \tau_p$, all exponentials have negative arguments, and the risk of overflow is avoided.

As can be seen, there is also a risk that the denominators $1 - \mu/\mu_0$ and $1 - k_{jn}\mu$ could be close to zero. However, this risk can be entirely eliminated by noting that when they are close to zero, there is in fact an exponential with argument close to zero in an integral in the preceding step. An exponential with zero argument is a constant, and the corresponding antiderivative does not have this denominator at all. Thus, if a denominator is close to zero, the corresponding term in the interpolation formulas is simply substituted with a term found by integrating the corresponding exponential term with zero argument. This can in fact be seen as an application of l'Hôpital's rules.

In the interpolation formulas everything is known except $I_0^-(\tau_0,\mu)$ and $I_L^+(\tau_L,\mu)$. $I_0^-(\tau_0,\mu)$ can be determined from the incident intensity at the top boundary. Then

$I_L^-(\tau_L, \mu)$ is calculated from the interpolation formulas, and using the boundary conditions $I_L^+(\tau_L, \mu)$ can be found. The interpolation formulas for the intensity give exactly the same result at the quadrature points as the discrete solution (3.17). They also satisfy the boundary conditions for all μ, albeit such conditions were imposed through (3.19) only at the quadrature points.

3.9. The δ-N Method. If the scattering is strongly forward-peaked, an accurate expansion of the phase function needs a large number, up to several hundreds or thousands, of terms. To maintain accuracy throughout the solution, a comparable number of terms are needed in the numerical quadrature used to approximate the integrals. This quickly gives very large eigenvalue problems and systems of equations, and since the computation time for these grows approximately as the third power of the size, the problem soon becomes intractable. The memory requirements also increase rapidly. To avoid this, a transformation proposed by Wiscombe [25], the δ-N method, can be applied to give a problem with a less peaked phase function.

The idea is to consider the beams scattered through the small angles within the sharp forward peak as unscattered, and truncate this peak from the phase function. The phase function is separated into the sum of a Dirac delta function in the forward direction and a truncated phase function, which is expanded in a series of Legendre polynomials with a much smaller number of terms, preferably equal to the number of quadrature points, i.e., $2N$.

On one hand, the phase function is directly expanded in Legendre polynomials as

$$p(\cos\Theta) = \sum_{l=0}^{N_{\text{large}}} (2l+1)\chi_l P_l(\cos\Theta).$$

On the other hand, the delta peak is first removed and then the remainder is expanded as

$$\begin{aligned} p(\cos\Theta) &= fp''(\cos\Theta) + (1-f)p'(\cos\Theta) \\ &\approx f\delta(1-\cos\Theta) + (1-f)\sum_{l=0}^{2N-1}(2l+1)\hat{\chi}_l P_l(\cos\Theta) \\ &\equiv \hat{p}_{\delta\text{-}N}(\cos\Theta), \end{aligned}$$

where f is a dimensionless parameter between 0 and 1 (f thus denotes the fraction of the phase function that is contained in the separated delta peak). Demanding that the coefficients for Legendre polynomial expansion are the same for p and $\hat{p}_{\delta\text{-}N}$, as long as they have common terms, yields

$$\chi_l = f + (1-f)\hat{\chi}_l, \quad \text{or} \quad \hat{\chi}_l = \frac{\chi_l - f}{1-f},\ l = 0, \ldots, 2N-1.$$

The expansion for $\hat{p}_{\delta\text{-}N}$ is truncated by demanding $\hat{\chi}_{2N} = 0$, which gives $f = \chi_{2N}$. Replacing p with $\hat{p}_{\delta\text{-}N}$ in the equation of radiative transfer and introducing $\tau' = (1-af)\tau$ and $a' = \frac{1-f}{1-af}a$ yields a structurally equivalent equation. Hence, the δ-N method does not change the mathematical form of the radiative transfer equation. It only changes the optical properties of the medium to make it appear less anisotropic.

Thus, the δ-N method allows handling of strongly forward-peaked phase functions (g close to 1) with maintained accuracy without a tremendously increased computational burden. The δ-N method also provides maintained accuracy for all g for

significantly lower N than otherwise needed. However, the closer g is to zero, the smaller N is needed anyway, so the savings in computation time diminish with decreasing $|g|$. The overhead introduced by the method is insignificant compared to the core calculations.

Morel [26] reports an alternative way of dealing with strongly forward-peaked scattering. He points out that it is not the accuracy of the truncated phase function expansion that matters, but rather the accuracy of the representation of the source function. Thus, if the solution is well represented by the given Legendre polynomial expansion, an accurate solution will be obtained regardless of the convergence of the truncated phase function expansion. Morel presents a Galerkin quadrature approach, which under this assumption treats the scattering exactly, and thus leaves the solution invariant to the δ-N transformation. Unfortunately, this approach is limited to the azimuthally averaged ($m = 0$) case, since the solution for Fourier components $m > 0$ cannot be well represented by Legendre polynomials, but needs the associated Legendre functions.

3.10. Intensity Correction Procedures. The accuracy of the intensity computation is generally improved by the use of the δ-N method except in the direction of the forward peak, but the δ-N method also introduces minor errors in other directions. However, combining the δ-N method with exact computation of low orders of scattering can considerably reduce the error. The purpose is to achieve high accuracy with small N, to speed up calculations. The TMS and IMS methods of Nakajima and Tanaka [27] serve to correct for single scattering and secondary and higher orders of scattering, respectively.

The phase function resulting from the δ-N method oscillates around the original phase function with a magnitude depending on the parameter f. This gives the computed intensities an oscillating behavior, which becomes more apparent the more peaked the phase function is. Since single scattering resembles the phase function, it would be a good idea to compute the single scattering exactly to account for errors due to the δ-N method.

Exact solutions for the single-scattered intensity are easy to derive. Using the integrodifferential equations for the Fourier components (3.6) without the multiple scattering terms, and allowing for the optical properties to vary between layers, gives elementary first-order differential equations that are readily solved.

The TMS method subtracts the erroneous single-scattered intensity obtained by using the scaled τ', the scaled a', and the phase function

$$p'(\cos\Theta) = \sum_{l=0}^{2N-1} (2l+1)\hat{\chi}_l P_l(\cos\Theta)$$

from the δ-N method, and adds back the exactly calculated single-scattered intensity obtained by using the scaled τ', $\frac{a}{1-af}$ (where the denominator is a consequence of the scaled τ'), and the exact phase function

$$p(\cos\Theta) = \sum_{l=0}^{N_{\text{large}}} (2l+1)\chi_l P_l(\cos\Theta)$$

with all available terms. This can be denoted $I_{TMS} = I' + \Delta I_{TMS} = I' - I'_{ss} + I^{corr}_{ss}$, where I' is the intensity computed by using the δ-N method, and I'_{ss} and I^{corr}_{ss} are the single-scattered intensities described above.

The TMS method gives a substantial improvement for the computed intensity, and the oscillations are suppressed. An error remains only in the direction of the forward peak. This is corrected in the IMS method by accounting for secondary and higher orders of scattering. Of course, exact solutions cannot be found for these corrections, since that would mean actually solving the overall problem. Instead, an exact solution can be derived symbolically, and then intelligent approximations need to be made in order to make the solution possible to use in the IMS method in practice. Reaching the final expression for the IMS method requires a substantial amount of algebra, and the original paper is also rather brief. However, the essentials will be outlined here.

The IMS method corrects only the intensity inside a cone centered on the forward peak direction, and thus affects only the downward intensity. Therefore, in this section, some simplifying notation can be used. All intensity variables I implicitly mean $I(\tau, -\mu, \varphi)$ and all angular integrals $\frac{1}{4\pi}\int_{4\pi} p \cdot I d\omega'$ implicitly mean

$$\frac{1}{4\pi}\int_0^{2\pi}\int_{-1}^{1} p(\tau, \mu', \varphi'; -\mu, \varphi) I(\tau, \mu', \varphi') d\mu' d\varphi'.$$

The optical properties, a, f, and $p \cdot I_{0b}$ implicitly mean, respectively, $a(\tau)$, $f(\tau)$, and $p(\tau, -\mu_0, \varphi_0; -\mu, \varphi) I_{0b}$.

Using the notation $I_{true} = I_{TMS} - \Delta I_{IMS}$, where I_{true} is the solution to the exact radiative transfer equation, what is left to be found is an expression for the IMS correction term $\Delta I_{IMS} = I_{TMS} - I_{true}$. Differentiating this, using the definitions of τ', I_{TMS}, I_{true}, and $p'' = \frac{1}{f}(p - (1-f)p')$, defining the δ-N multiple-scattered intensity as $I'_{mult} = I' - I'_{ss}$, and algebraically rearranging gives

$$-\mu \frac{d}{d\tau}(\Delta I_{IMS}) = \Delta I_{IMS} - \frac{a}{4\pi}\int_{4\pi} p \cdot \Delta I_{IMS} d\omega' - (Q_1 + Q_2 + Q_3), \tag{3.28}$$

where

$$Q_1 = af\left(I'_{mult} - \frac{1}{4\pi}\int_{4\pi} p'' \cdot I'_{mult} d\omega'\right),$$
$$Q_2 = af\left(I^{corr}_{ss} - \frac{1}{4\pi}\int_{4\pi} p'' \cdot I^{corr}_{ss} d\omega'\right),$$

and

$$Q_3 = \frac{a}{4\pi} p \cdot I_{0b}\left(e^{-\tau'/\mu_0} - e^{-\tau/\mu_0}\right) - \frac{a}{4\pi}(1-f)\int_{4\pi} p' \cdot (I^{corr}_{ss} - I'_{ss})\, d\omega'.$$

This exact equation for the IMS correction term ΔI_{IMS} is more complicated than the original radiative transfer equation, so several approximations need to be made in order to make the IMS method practically useful. First,

$$-\frac{a}{4\pi}\int_{4\pi} p \cdot \Delta I_{IMS} d\omega' \approx 0$$

and $Q_1 \approx 0$, since their contribution to the narrow forward peak, where the IMS correction method is used, is negligible. Second,

$$Q_2 \approx \frac{I_{0b}}{4\pi}\frac{(af)^2}{1-af}\frac{e^{-\tau'/\mu_0}}{\mu_0}\tau'\left(p'' - \frac{1}{4\pi}\int_{4\pi} p'' \cdot p'' d\omega'\right)$$

and

$$Q_3 \approx \frac{I_{0b}}{4\pi}\frac{(af)^2}{1-af}\frac{e^{-\tau'/\mu_0}}{\mu_0}\tau' p'',$$

where the reasons are more complex, so the interested reader is directed to the original paper by Nakajima and Tanaka [27]. Finally, the IMS method uses vertically averaged optical properties:

$$\bar{a} = \left(\sum_{n=1}^{p} a_n\tau_n\right) \Big/ \left(\sum_{n=1}^{p} \tau_n\right),$$

$$\bar{f} = \left(\sum_{n=1}^{p} f_n a_n\tau_n\right) \Big/ \left(\sum_{n=1}^{p} a_n\tau_n\right),$$

$$\chi'_{l,n} = \begin{cases} f_n, & l \le 2N-1, \\ \chi_{l,n}, & l > 2N-1, \end{cases}$$

$$\bar{\chi}_l = \left(\sum_{n=1}^{p} \chi'_{l,n} a_n\tau_n\right) \Big/ \left(\sum_{n=1}^{p} f_n a_n\tau_n\right),$$

$$p''(\cos\Theta) = \sum_{l=0}^{N_{\text{large}}} (2l+1)\bar{\chi}_l P_l(\cos\Theta),$$

$$\mu'_0 \equiv \frac{1}{1-\bar{a}\bar{f}}\mu_0,$$

where n is the layer index.

Equation (3.28) for the IMS correction term ΔI_{IMS} then becomes a first-order differential equation that can be solved by integrating from 0 to τ, using $e^{\tau/\mu}$ as an integrating factor. This gives the IMS correction term, which is then expanded into a Fourier cosine series to provide the final expression that is used in the IMS method.

Thus a single Fourier component becomes

$$\Delta I^m_{IMS} = \frac{I_{0b}}{4\pi}\frac{(\bar{a}\bar{f})^2}{1-\bar{a}\bar{f}}(2-\delta_{0m})p^m_{IMS}(-\mu'_0,-\mu)\frac{e^{-\tau/\mu}}{\mu\mu'_0}\int_0^\tau e^{(1/\mu-1/\mu'_0)t}t\,dt, \tag{3.29}$$

where

$$p^m_{IMS}(-\mu'_0,-\mu) = \sum_{l=m}^{N_{\text{large}}} (2l+1)(2\bar{\chi}_l - \bar{\chi}^2_l)\Lambda^m_l(-\mu'_0)\Lambda^m_l(-\mu).$$

The intensity correction procedures further enhance the handling of strongly forward-peaked phase functions beyond the capabilities of the δ-N method. These procedures also give maintained accuracy for all g for significantly lower N than otherwise needed. However, the closer g is to zero, the smaller N is needed anyway, so at some point the possible savings in computation time are smaller than the overhead introduced by the correction procedures. The correction procedures should therefore not be used in those cases. The additional time taken for the intensity correction procedures consists of evaluating Legendre functions Λ^m_l for the larger l and m that are used.

3.11. Computational Shortcuts. As shown by King [28], the azimuthal dependence of the intensity typically converges well before the loop over Fourier components has ended. Since it is the outermost loop, much is gained if it can be terminated earlier. It is therefore beneficial to break the azimuthal loop when a convergence criterion has been met, for example, when the quotient of the absolute value of a Fourier component and the cumulative sum of components is smaller than a given limit. This saves a significant amount of computation time in the vast majority of cases.

There is an obvious computational shortcut that allows for much faster calculation of variables that depend only on the azimuthally averaged intensity, which is given by the zeroth Fourier component. Among these variables are total reflectance, total transmittance, total absorptance, and flux. When such variables are all that is required, the azimuthal loop is broken after the first time instead of fulfilling the prescribed $2N$ times, thus giving a significant reduction in computation time.

4. Implementation and Performance. The solution method described in this paper has recently been implemented in MATLAB under the name DORT2002, and it is now being used in the paper and printing industries for light scattering simulations. In the current process of replacing an older generation of simulation tools in these sectors of industry, there is a need for more accurate models. These will offer more understanding and deeper insight in the processes of light scattering in such complex media as paper. The effect of the different paper constituents on light scattering may then be investigated theoretically as well, with higher accuracy than real measurements. Application areas for DORT2002 therefore include theoretical model comparisons, but also fine-tuning the papermaking process, designing new paper qualities, color management from prepress to print, and evaluation of printing techniques. DORT2002 is, however, also intended as a general tool for radiative transfer problems, and it can be obtained from the author for evaluation.

Performance and application of DORT2002 have been studied in an extensive test series, which will be reported elsewhere. However, it may be appropriate to give a short summary here. Tests show that the preconditioner for the system of equations corresponding to the boundary and continuity conditions works very well, giving a condition number close to 1 in most cases, and around 30 in the worst case. They also show that the problem is very ill-conditioned without the preconditioner, having a condition number near the largest positive floating-point number for the system. It is also shown that the reduced eigenvalue problem is very well-conditioned, giving a condition number close to 1 in all cases. DORT2002 is shown to converge when N increases.

Performance tests show that the steps that are taken to improve the stability and speed of DORT2002 are very successful, together giving an unconditionally stable solution procedure to a problem previously considered numerically intractable, and together decreasing computation time with a factor of 1,000–10,000 in typical cases and with a factor up to and beyond 10,000,000 in extreme cases. Application tests show very good agreement of DORT2002 with three established models of different types and implementations when applied to different sets of relevant test problems, which gives strong support for the accuracy of DORT2002.

5. Open Questions and Future Work. The TMS and IMS methods take relatively little time in themselves, but far more time is taken by the evaluation of the normalized associated Legendre functions, $\Lambda_l^m(u)$, for the larger l and m needed for these methods. Any studies that result in faster ways of evaluation of $\Lambda_l^m(u)$ for large l and m would be welcome.

One bottleneck that remains is the generation in MATLAB of the sparse matrix in the system of equations corresponding to the boundary and continuity conditions. Although the values and the indices of the nonzero elements are known, the assigning of these values to the sparse matrix is unsatisfactorily time consuming in MATLAB, to the extent that this purely administrative part of the code consumes a significant part of the execution time. Since all computational parts of the code are already so optimized, this item is the first candidate for improving the speed of the code. This problem remains, although the current implementation has been worked out in cooperation with leading MATLAB experts and although the implementation, to the author's knowledge, is the best that can be done in MATLAB today. Improvements in this direction could well be considered in future versions of MATLAB.

As an upcoming research activity, the inverse problem for the model presented in this paper will be studied. This includes the study and development of fast and numerically stable algorithms for parameter estimation. The parameter estimation will be carried out to fit model simulations to angle-resolved light scattering measurements or to desired angle-resolved light scattering patterns. This opens the possibility of indirectly measuring parameters that are hard to determine in other ways, but also to constructing materials with designed optical properties. The starting point is known intensities in different directions in the form of measurements or design goals, and known boundary conditions. In the simplest case the single scattering albedo, a, and the asymmetry factor, g, are estimated. More complicated cases are multilayer structures where a and g are estimated for all layers, possibly with different values in every layer. In addition to this, it will be necessary to deal with the problem of surface effects such as gloss in real-life measurements.

Acknowledgment. The author wishes to thank two anonymous referees for their valuable comments on the manuscript, and especially for pointing out some explicit references in the neutron transport and nuclear engineering areas.

REFERENCES

[1] A. Schuster, *Radiation through a foggy atmosphere*, Astrophys. J., 21(1905), pp. 1–22. (Reprinted in *Selected Papers on the Transfer of Radiation*, D. H. Menzel, Dover, New York, 1966, pp. 3–24.)

[2] P. Kubelka and F. Munk, *Ein Beitrag zur Optik der Farbanstriche*, Z. Tech. Phys., 11a (1931), pp. 593–601.

[3] P. Kubelka, *New contributions to the optics of intensely light-scattering materials. Part* I, J. Opt. Soc. Amer., 38 (1948), pp. 448–457.

[4] P. Kubelka, *New contributions to the optics of intensely light-scattering materials. Part* II, J. Opt. Soc. Amer., 44 (1954), pp. 330–335.

[5] G. C. Wick, *Über ebene Diffusionsprobleme*, Z. Phys. 120 (1943), pp. 702–718.

[6] S. Chandrasekhar, *On the radiative equilibrium of a stellar atmosphere*, Astrophys. J., 99 (1944), pp. 180–190.

[7] S. Chandrasekhar, *On the radiative equilibrium of a stellar atmosphere* II, Astrophys. J., 100 (1944), pp. 76–86.

[8] S. Chandrasekhar, *Radiative Transfer*, Dover, New York, 1960.

[9] P. S. Mudgett and L. W. Richards, *Multiple scattering calculations for technology*, Appl. Opt., 10 (1971), pp. 1485–1502.

[10] P. S. Mudgett and L. W. Richards, *Multiple scattering calculations for technology* II, J. Colloid Interf. Sci., 39 (1972), pp. 551–567.

[11] Lord Rayleigh, *On the light from the sky, its polarization and colour*, Philos. Mag., 41 (1871), pp. 107–120, 274–279. (Reprinted in Scientific Papers by Lord Rayleigh, Vol. I: 1869–1881, No. 8, Dover, New York, 1964.)

[12] G. Mie, *Beiträge zur Optik trüber Medien, Speziell Kolloidaler Metallösungen*, Ann. Phys., 25 (1908), pp. 377–445.

[13] L. G. Henyey and J. L. Greenstein, *Diffuse radiation in the galaxy*, Astrophys. J., 93 (1941), pp. 70–83.
[14] W. Magnus and F. Oberhettinger, *Formulas and Theorems for the Functions of Mathematical Physics*, Chelsea, New York, 1949.
[15] G. H. Golub and J. H. Welsch, *Calculation of Gauss quadrature rules*, Math. Comp., 23 (1969), pp. 221–230.
[16] J. B. Sykes, *Approximate integration of the equation of transfer*, Monthly Not. Roy. Astr. Soc., 111 (1951), pp. 377–386.
[17] K. Stamnes and R. A. Swanson, *A new look at the discrete ordinate method for radiative transfer calculations in anisotropically scattering atmospheres*, J. Atmos. Sci., 38 (1981), pp. 387–399.
[18] Z. Jin and K. Stamnes, *Radiative transfer in nonuniformly refracting media such as the atmosphere/ocean system*, Appl. Opt., 33 (1994), pp. 431–442.
[19] R. C. De Barros and E. W. Larsen, *A numerical method for one-group slab-geometry discrete ordinates problems with no spatial truncation error*, Nucl. Sci. Eng., 104 (1990), pp. 199–208.
[20] R. C. De Barros and E. W. Larsen, *A spectral nodal method for one-group x,y-geometry discrete ordinates problems*, Nucl. Sci. Eng., 111 (1992), pp. 34–45.
[21] R. C. De Barros, F. C. da Silva, and H. A. Filho, *Recent advances in spectral nodal methods for x,y-geometry discrete ordinates deep penetration and eigenvalue problems*, Progress in Nuclear Energy, 35 (1999), pp. 293–331.
[22] A. Badruzzaman, *Nodal Methods in Transport Theory*, Advances in Nuclear Science and Technology 21, J. Lewins and M. Becker, eds., Plenum Press, New York, 1990.
[23] K. Stamnes and P. Conklin, *A new multi-layer discrete ordinate approach to radiative transfer in vertically inhomogeneous atmospheres*, J. Quant. Spectrosc. Radiat. Transfer, 31 (1984), pp. 273–282.
[24] K. Stamnes, *On the computation of angular distributions of radiation in planetary atmospheres*, J. Quant. Spectrosc. Radiat. Transfer, 28 (1982), pp. 47–51.
[25] W. J. Wiscombe, *The delta-M method: Rapid yet accurate radiative flux calculations for strongly asymmetric phase functions*, J. Atmos. Sci., 34 (1977), pp. 1408–1422.
[26] J. E. Morel, *A hybrid collocation-Galerkin-S_n method for solving the Boltzmann transport equation*, Nucl. Sci. Eng., 101 (1989), pp. 72–87.
[27] T. Nakajima and M. Tanaka, *Algorithms for radiative intensity calculations in moderately thick atmospheres using a truncation approach*, J. Quant. Spectrosc. Radiat. Transfer, 40 (1988), pp. 51–69.
[28] M. D. King, *Number of terms required in the Fourier expansion of the reflection function for optically thick atmospheres*, J. Quant. Spectrosc. Radiat. Transfer, 30 (1983), pp. 143–161.
[29] G. H. Golub and C. F. Van Loan, *Matrix Computations*, 3rd ed., Johns Hopkins University Press, Baltimore, MD, 1996.