PERGAMON

Journal of Quantitative Spectroscopy & Radiative Transfer 72 (2002) 789–802

Journal of Quantitative Spectroscopy & Radiative Transfer

www.elsevier.com/locate/jqsrt

Radiative perturbation theory for polarized radiances: 1. Theory

Yalong Tian, Michael A. Box [*]

School of Physics, University of New South Wales, Sydney, NSW 2052, Australia

Received 8 January 2001; accepted 9 May 2001

Abstract

Radiative perturbation theory has proven to be a useful tool in radiative transfer calculations, especially in situations where repeated solution of the radiative transfer equation is required. So far however, its use has been restricted to non-polarized situations, including such applications as surface fluxes, UV indices, and the inversion of satellite radiance observations. Here, we extend the structure of radiative perturbation theory to incorporate the full Stokes formalism of polarization, to obtain the relevant equations for the first order term. This formalism will be applied to fluxes in a follow-up paper, and eventually to satellite observations. © 2002 Elsevier Science Ltd. All rights reserved.

Keywords: Radiative transfer; Polarization

1. Introduction

Radiative perturbation theory has, over the years, shown itself to be a valuable tool in performing radiative transfer calculations. With this technique, one first performs a detailed radiative transfer computation for a suitably chosen ‘base case’ atmospheric (optical) model, and then perturbs about this base case by evaluating some (relatively) simple numerical integrals [1,2]. Like all forms of (first order) perturbation, the accuracy of the perturbation prediction, relative to the true value as obtained via a full radiative transfer calculation, decreases with the distance (in parameter space) one moves from the original model. Nevertheless, we have been able to obtain some very impressive results by a suitable (non-linear) formulation of the technique [3,4].

The main applications of radiative perturbation theory are, of course, in situations where repeated solution of the radiative transfer equation for closely related atmospheric models is

---

[*] Corresponding author. Tel.: +612-9385-4545; fax: +612-9385-6060.
E-mail address: mab@newt.phys.unsw.edu.au (M.A. Box).

PII: S0022-4073(01)00158-3

needed. Examples of this include the computation of UV indices (which requires the integration of the product of the surface flux, and a biological susceptibility, over wavelength [5,6]), and retrieval algorithms requiring iteration of the radiative transfer equation [7,8].

So far, all applications of radiative perturbation theory have been confined to the scalar, or non-polarized version of the radiative transfer equation, as the number of situations where the vector, or polarized form is required is relatively small. However, there is at least one situation where application of the perturbation technique to polarized radiation is of interest, and that is the retrieval of atmospheric geophysical data from observations of polarized radiation. Such data is (potentially) available either from a Cimel radiometer [9] at the Earth's surface, or the orbiting POLDER instrument [10].

While a number of algorithms have been implemented to analyse such data, including both table look-up and iteration, we believe that radiative perturbation theory may well prove to be a very valuable addition to this collection. Application of perturbation theory to radiances (whether polarized or not) involves a number of steps beyond its application to fluxes [7,8]. For this reason, we decided to embark on this challenge as a two step process. The first step, as contained in this paper, is to develop the general formalism of radiative perturbation theory for polarized radiation. In Section 2 we provide a short outline of radiative perturbation theory, so that its formalism may be extended in Sections 3 and 4 to the polarized case. Possible applications are outlined in Section 5. In further papers we will apply this formalism firstly to fluxes, and later to satellite observations.

2. Radiative perturbation theory

2.1. Operator formulation

The formulation of the radiative perturbation technique is best handled using an operator notation [11]. We start by writing the radiative transfer equation in operator form as

$$LI = Q, \tag{1}$$

where I is the radiance field (distribution function), Q the source of radiation (the solar beam, for example), and L is the transport operator [11], which may be written as

$$L \equiv \mu\frac{\partial}{\partial z} + \sigma_{\mathrm{t}} - \sigma_{\mathrm{s}} \int_{4\pi} \mathrm{d}\Omega' p(z, \Omega' \to \Omega) \circ . \tag{2}$$

Here $\sigma_{\mathrm{t}}(z)$, $\sigma_{\mathrm{s}}(z)$ and p are the vertical profiles of the extinction cross section, scattering cross section (per unit volume), and the phase function, respectively. The notation $\circ$ is used to denote an integral operator, not a definite integral. Once the radiative transfer equation has been solved, and we have (at least in principle) the radiance field, $I(z, \Omega)$, we may extract any information we choose—we refer to this as a 'radiative effect', E—by the use of a suitable 'response function', R:

$$E = \langle I, R\rangle, \tag{3}$$

where angular brackets denote integration over the phase space variables:

$$\langle f, g \rangle \equiv \iint f(z,\Omega) g(z,\Omega)\,\mathrm{d}z\,\mathrm{d}\Omega. \tag{4}$$

In addition, it is also necessary to introduce the adjoint version of the radiative transfer equation:

$$L^+ I^+ = Q^+. \tag{5}$$

Here Q^+ is the adjoint source (essentially arbitrary at this stage), I^+ is the adjoint radiance (the solution of the adjoint transfer equation), and L^+ is the adjoint transport operator. It may be shown [11] that L^+ is identical to L, except for a change of sign in the first term (the derivative with respect to z). Solving the adjoint transport equation is actually no more difficult, in general, than solving the ordinary transport equation, and sometimes easier, as we may often avoid the complications of a strongly collimated solar beam.

If we now decide to use the response function, R, as the adjoint source in this equation, then it may be shown [11] using the definition of an adjoint operator, that we may also obtain our desired effect, E, via

$$E = \langle I^+, Q \rangle. \tag{6}$$

Effectively we have two paths to any desired effect. The normal (or forward) way is to start with the appropriate source, solve the radiative transfer equation, and then extract the effect via the appropriate response function. (This last step is usually trivial, of course.) The adjoint way is to start with the response function as an adjoint source, solve the adjoint transport equation, and then extract the effect via the original source function—essentially working backwards.

2.2. Perturbation

Assume now that we have solved *both* the radiative transfer equation, *and* the adjoint equation, for a certain source and effect, in the case of an atmospheric optical model characterized by a transport operator L_0 (the base case). Thus we have the two radiance fields I_0 and I_0^+, and from either may obtain (the value of) the desired effect, E_0. We now ask the simple question: how much would the value of E change if we were to make a (relatively) small perturbation to the atmospheric optical model? The answer, to first order, is [1]

$$\Delta E = -\langle I_0^+, \Delta L I_0 \rangle, \tag{7}$$

where ΔL is the difference in the two atmospheric models, as reflected in the difference between their corresponding transport operators. In effect, what we are doing is performing a Taylor expansion (to first order) of (the value of) our effect with respect to certain of the model parameters. Simple examples of such perturbation include an increase or decrease in aerosol loading, or a change in aerosol optical properties [2–4].

3. Operator formalism for polarization

In order to perform perturbation calculations in the case of polarized radiative transfer, we must formulate the problem in operator notation. This will be a two step process. Firstly we must write the transport equation in operator form, and then we will be able to write the perturbation expressions. However, as we will see, this involves a number of choices.

The standard formulation of polarized radiative transfer involves an equation which is apparently the same as the non-polarized, or scalar, equation, provided we replace the intensity field by a four-component Stokes vector field, and replace the scalar phase function by a 4×4 phase matrix:

$$\mu \frac{\partial \vec{I}}{\partial z} + \sigma_t \vec{I} - \sigma_s \int \mathrm{d}\Omega' \, \mathbf{P}(\Omega' \to \Omega) \vec{I} = \vec{Q}. \tag{8}$$

In writing this equation we have introduced the notations that a superposed arrow denotes a four component vector, while a bold symbol denotes a 4×4 matrix. The intensity field is usually represented by the Stokes vector:

$$\vec{I} \equiv \begin{pmatrix} I_l \\ I_r \\ U \\ V \end{pmatrix}. \tag{9}$$

In the case of scattering of solar radiation, the source vector becomes

$$\vec{Q} \equiv \vec{F}_0 = \begin{pmatrix} 1/2 \\ 1/2 \\ 0 \\ 0 \end{pmatrix} F_0, \tag{10}$$

where F_0 is the solar constant. (In practice, when solving this equation, one normally separates the diffuse intensity from the direct beam.)

We are now in a position to write the polarized transport equation in operator form:

$$\mathbf{L} \vec{I} = \vec{Q}, \tag{11}$$

where $\mathbf{L}$ is a 4×4 matrix operator, generalized from the scalar case:

$$\mathbf{L} = \mu \frac{\partial}{\partial z} \mathbf{M} + \sigma_t \mathbf{M} - \sigma_s \int \mathrm{d}\Omega' \, \mathbf{P}(\Omega' \to \Omega) \circ . \tag{12}$$

Here, $\mathbf{M}$ is the four-by-four identity matrix. In addition, we will need the adjoint equation, and the adjoint transport operator:

$$\mathbf{L}^+ \vec{I}^+ = \vec{Q}^+. \tag{13}$$

This time the adjoint transport operator differs from the standard operator in both the sign of the derivative term, and the fact that $\mathbf{P}$ must be replaced by its transpose in order to ensure that this operator satisfies the basic property of an adjoint (Appendix A.1), namely

$$\langle \mathbf{L}^+ \vec{I}^+, \vec{I} \rangle = \langle \vec{I}^+, \mathbf{L} \vec{I} \rangle. \tag{14}$$

In writing this equation we understand that the first vector quantity in each case is first transposed, and then an inner product taken with the second vector quantity, yielding a scalar quantity, which is then integrated over the phase space variables:

$$\langle \vec{f}, \vec{g} \rangle \equiv \int\!\!\int \mathrm{d}z\, \mathrm{d}\Omega\, \tilde{\vec{f}} \vec{g}, \tag{15}$$

where the tilde, $\sim$, denotes the transpose of a vector or matrix.

Note that it is also necessary to impose adjoint boundary conditions similar to those imposed in the scalar case, that is

$$\begin{aligned} \vec{I}(z_{\mathrm{T}}, \mu, \phi) &= \vec{0}, \quad -1 \leqslant \mu < 0, \\ \vec{I}(0, \mu, \phi) &= \vec{0}, \quad 0 < \mu \leqslant 1, \\ \vec{I}^{+}(z_{\mathrm{T}}, \mu, \phi) &= \vec{0}, \quad 0 < \mu \leqslant 1, \\ \vec{I}^{+}(0, \mu, \phi) &= \vec{0}, \quad -1 \leqslant \mu < 0, \end{aligned} \tag{16}$$

where z_{T} denotes the top of the atmosphere. That is to say, we have no incoming radiances, and no outgoing adjoint radiances.

It is straightforward to show (Appendix A.2) that solution of the vector adjoint transport equation follows essentially the same route as solution of the scalar adjoint. That is, one first solves for the 'pseudo-radiance' vector produced by direction reversing within the adjoint source (the response function), and then reverses directions to convert pseudo-radiance to actual adjoint radiance information.

The effects which we are ultimately seeking, such as the intensity and polarization components as measured by a suitable instrument, are inherently vector quantities (which may or may not be measured independently). Even the flux at the surface can be considered as having two (Stokes) components (see below). This adds a new complication into our definitions and notations. We start by defining the modified inner product

$$\{\vec{f}, \vec{g}\} \equiv \int\!\!\int \mathrm{d}z\, \mathrm{d}\Omega\, \mathbf{f} \vec{g}, \tag{17}$$

where the matrix, $\mathbf{f}$, is derived from the corresponding vector according to

$$\mathbf{f} = \begin{pmatrix} f_1 & 0 & 0 & 0 \\ 0 & f_2 & 0 & 0 \\ 0 & 0 & f_3 & 0 \\ 0 & 0 & 0 & f_4 \end{pmatrix}. \tag{18}$$

In order to obtain an effect vector, (E_1, E_2, E_3, E_4), we need the corresponding response function, also as a vector. For example, if we are interested in the flux at an altitude z_0, then

$$\vec{R} = \mu \delta(z - z_0) \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix} \tag{19}$$

(If only the upward or downward component is required, than we may include the Heaviside step function into this definition.) Once we have solved the (polarized) transport equation, we may extract (the components of) our effect vector via

$$\vec{E} = \{\vec{R}, \vec{I}\}. \tag{20}$$

Alternatively, we could solve the adjoint equation (above), and then use

$$\vec{E} = \{\vec{I}^+, \vec{Q}\}. \tag{21}$$

For a number of reasons, including notational convenience, it will prove useful to be able to write expressions for the individual components of the effect vector. We start by defining a set of 4×4 matrix operators, $\mathbf{L}_i$ and $\mathbf{L}_i^+$, which are, respectively, the ith row and ith column of the full operators. For example

$$\mathbf{L}_1 = \begin{pmatrix} L_{11} & 0 & 0 & 0 \\ L_{21} & 0 & 0 & 0 \\ L_{31} & 0 & 0 & 0 \\ L_{41} & 0 & 0 & 0 \end{pmatrix} \tag{22}$$

and

$$\mathbf{L}_1^+ = \begin{pmatrix} L_{11}^+ & L_{12}^+ & L_{13}^+ & L_{14}^+ \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}. \tag{23}$$

With these definitions, we may write down the twin expressions from which we may obtain the value of a single component of the effect vector:

$$E_i = \langle \vec{I}^+, \mathbf{L}_i \vec{I} \rangle = \langle \mathbf{L}_i^+ \vec{I}^+, \vec{I} \rangle. \tag{24}$$

4. Perturbation theory for polarization

Let us now consider two atmospheric optical models, characterised by two different transport operators, $\mathbf{L}_0$ and $\mathbf{L}$, with their corresponding intensity (vector) fields, and as a consequence, their respective (values of) a certain radiative effect. We may express the connection between these two cases in perturbation notation as follows:

$$\vec{I} = \vec{I}_0 + \Delta \vec{I}, \tag{25a}$$

$$\mathbf{L} = \mathbf{L}_0 + \Delta \mathbf{L}, \tag{25b}$$

$$\vec{E} = \vec{E}_0 + \Delta \vec{E}. \tag{25c}$$

In the case of the transport (matrix) operator, and the effect vector, we may also define the perturbations for the individual components, as necessary. From these definitions, we may show (Appendix A.3) that

$$\Delta E_i \cong -\langle \vec{I}_0^+, \Delta \mathbf{L}_i \vec{I}_0 \rangle, \tag{26}$$

where this approximate equality is to be understood in terms of first order perturbation.

We turn now to the general expressions for the perturbation of the components of a given effect vector. We start by expanding the intensity Stokes vector in Fourier series (a similar expansion will also apply to the adjoint).

$$\vec{I}(z,\mu,\phi) = \sum_{n=0}^{N} \begin{pmatrix} I_1^n(z,\mu)\cos n(\phi_0-\phi) \\ I_2^n(z,\mu)\cos n(\phi_0-\phi) \\ I_3^n(z,\mu)\sin n(\phi_0-\phi) \\ I_4^n(z,\mu)\sin n(\phi_0-\phi) \end{pmatrix}. \tag{27}$$

The Rayleigh scattering phase matrix has been given explicitly by Chandrasekhar [12] and Coulson [13]. The Mie scattering phase matrix depends, of course, on the particular aerosols or clouds being considered. Nevertheless, Dave [14] showed that, in general it will have the following Fourier expansion properties:

$$P_{ij}[\mu,\mu',(\phi'-\phi)] = \sum_{n=0}^{N} P_{ij}^n(\mu,\mu')\cos n(\phi'-\phi) \tag{28}$$

with

$$\begin{aligned} P_{ij}^n(-\mu,-\mu') &= P_{ji}^n(\mu,\mu'), \\ P_{ij}^n(\mu,-\mu') &= P_{ji}^n(-\mu,\mu') \end{aligned} \tag{29}$$

when $ij = 11, 12, 21, 22, 33, 34, 43$ and 44, while

$$P_{ij}[\mu,\mu',(\phi'-\phi)] = \sum_{n=0}^{N} P_{ij}^n(\mu,\mu')\sin(\phi'-\phi) \tag{30}$$

with

$$\begin{aligned} P_{ij}^n(-\mu,-\mu') &= -P_{ji}^n(\mu,\mu'), \\ P_{ij}^n(\mu,-\mu') &= -P_{ji}^n(-\mu,\mu') \end{aligned} \tag{31}$$

when $ij = 13, 14, 23, 24, 31, 32, 41$ and 42.

Using these results, we may write for the perturbation transport operator, $\Delta\mathbf{L}$

$$\begin{aligned}\Delta\mathbf{L} ={}& (\sigma_{\rm t}(z) - \sigma_{\rm t}^0(z))\mathbf{M}\\
&- \sigma_{\rm s}(z)\sum_{m=0}^{N}\int_0^{2\pi}\int_{-1}^{1}\begin{pmatrix} \begin{pmatrix} P_{11}^m & P_{12}^m \\ P_{21}^m & P_{22}^m \end{pmatrix}\cos m(\phi'-\phi) & \begin{pmatrix} P_{13}^m & P_{14}^m \\ P_{23}^m & P_{24}^m \end{pmatrix}\sin m(\phi'-\phi) \\ \begin{pmatrix} P_{31}^m & P_{32}^m \\ P_{41}^m & P_{42}^m \end{pmatrix}\sin m(\phi'-\phi) & \begin{pmatrix} P_{33}^m & P_{34}^m \\ P_{43}^m & P_{44}^m \end{pmatrix}\cos m(\phi'-\phi) \end{pmatrix} \mathrm{d}\mu'\,\mathrm{d}\phi'\\
&+ \sigma_{\rm s}^0(z)\sum_{m=0}^{N}\int_0^{2\pi}\int_{-1}^{1}\begin{pmatrix} \begin{pmatrix} P_{011}^m & P_{012}^m \\ P_{021}^m & P_{022}^m \end{pmatrix}\cos m(\phi'-\phi) & \begin{pmatrix} P_{013}^m & P_{013}^m \\ P_{023}^m & P_{024}^m \end{pmatrix}\sin m(\phi'-\phi) \\ \begin{pmatrix} P_{031}^m & P_{032}^m \\ P_{041}^m & P_{042}^m \end{pmatrix}\sin m(\phi'-\phi) & \begin{pmatrix} P_{033}^m & P_{034}^m \\ P_{043}^m & P_{044}^m \end{pmatrix}\cos m(\phi'-\phi) \end{pmatrix} \mathrm{d}\mu'\,\mathrm{d}\phi'.\end{aligned} \tag{32}$$

Because of the complexity of the various expressions we need to play with, we will examine the perturbation of the components of the effect vector one at a time. Since the perturbation operator $\Delta\mathbf{L}_1$ will have the same column structure as $\mathbf{L}_1$, we see that we may write for the first component

$$\begin{aligned}\Delta E_1 ={}& -\int_0^{z_{\rm T}}\mathrm{d}z\int_{-1}^{1}\mathrm{d}\mu\int_0^{2\pi}\mathrm{d}\phi\left(\sum_{l=0}^{N}\begin{pmatrix} I_1^{l+}\cos l(\phi_0-\phi) \\ I_2^{l+}\cos l(\phi_0-\phi) \\ I_3^{l+}\sin l(\phi_0-\phi) \\ I_4^{l+}\sin l(\phi_0-\phi) \end{pmatrix}^{\rm T}\right.\\
&\times\left(\left(\begin{pmatrix} \sigma_{\rm t}(z)-\sigma_{\rm t}^0(z) & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \times\sum_{n=0}^{N}\begin{pmatrix} I_1^n(z,\mu)\cos n(\phi_0-\phi) \\ I_2^n(z,\mu)\cos n(\phi_0-\phi) \\ I_3^n(z,\mu)\sin n(\phi_0-\phi) \\ I_4^n(z,\mu)\sin n(\phi_0-\phi) \end{pmatrix}\right.\right.\\
&-\int_{-1}^{1}\mathrm{d}\mu'\int_0^{2\pi}\mathrm{d}\phi'\sum_{m=0}^{N}\begin{pmatrix} (\sigma_{\rm s}(z)P_{11}^m - \sigma_{\rm s}^0(z)P_{011}^m)\cos m(\phi'-\phi) & 0 & 0 & 0 \\ (\sigma_{\rm s}(z)P_{21}^m - \sigma_{\rm s}^0(z)P_{021}^m)\cos m(\phi'-\phi) & 0 & 0 & 0 \\ (\sigma_{\rm s}(z)P_{31}^m - \sigma_{\rm s}^0(z)P_{031}^m)\sin m(\phi'-\phi) & 0 & 0 & 0 \\ (\sigma_{\rm s}(z)P_{41}^m - \sigma_{\rm s}^0(z)P_{041}^m)\sin m(\phi'-\phi) & 0 & 0 & 0 \end{pmatrix}\end{aligned}$$

$$\times\sum_{n=0}^{N}\begin{pmatrix} I_1^n(z,\mu')\cos n(\phi_0-\phi') \\ I_2^n(z,\mu')\cos n(\phi_0-\phi') \\ I_3^n(z,\mu')\sin n(\phi_0-\phi') \\ I_4^n(z,\mu')\sin n(\phi_0-\phi') \end{pmatrix}\Bigg)\Bigg)\Bigg). \tag{33}$$

We may now make use of the orthogonality relations of the trig functions to perform all the azimuthal integrals to obtain the final result:

$$\begin{aligned}\Delta E_1 = -\pi\sum_{n=0}^{N}\int_0^{z_\mathrm{T}}\mathrm{d}z\int_{-1}^{1}\mathrm{d}\mu\Big\{&(1+\delta_{0n})(\sigma_\mathrm{t}(z)-\sigma_\mathrm{t}^0(z))I_1^{n+}(z,\mu)I_1^n(z,\mu)\\ &-\pi\Big[(1+\delta_{0n})^2 I_1^{n+}(z,\mu)\int_{-1}^{1}\mathrm{d}\mu'(\sigma_\mathrm{s}(z)P_{11}^n-\sigma_\mathrm{s}^0(z)P_{011}^n)I_1^n(z,\mu')\\ &+(1+\delta_{0n})^2 I_2^{n+}(z,\mu)\int_{-1}^{1}\mathrm{d}\mu'(\sigma_\mathrm{s}(z)P_{21}^n-\sigma_\mathrm{s}^0(z)P_{021}^n)I_1^n(z,\mu')\\ &+(1-\delta_{0n})I_3^{n+}(z,\mu)\int_{-1}^{1}\mathrm{d}\mu'(\sigma_\mathrm{s}(z)P_{31}^n-\sigma_\mathrm{s}^0(z)P_{031}^n)I_1^n(z,\mu')\\ &+(1-\delta_{0n})I_4^{n+}(z,\mu)\int_{-1}^{1}\mathrm{d}\mu'(\sigma_\mathrm{s}(z)P_{41}^n-\sigma_\mathrm{s}^0(z)P_{041}^n)I_1^n(z,\mu')\Big]\Big\}.\end{aligned} \tag{34}$$

In a similar fashion we find that

$$\begin{aligned}\Delta E_2 = -\pi\sum_{n=0}^{N}\int_0^{z_\mathrm{T}}\mathrm{d}z\int_{-1}^{1}\mathrm{d}\mu\Big\{&(1+\delta_{0n})(\sigma_\mathrm{t}(z)-\sigma_\mathrm{t}^0(z))I_2^{n+}(z,\mu)I_2^n(z,\mu)\\ &-\pi\Big[(1+\delta_{0n})^2 I_1^{n+}(z,\mu)\int_{-1}^{1}\mathrm{d}\mu'(\sigma_\mathrm{s}(z)P_{12}^n-\sigma_\mathrm{s}^0(z)P_{012}^n)I_2^n(z,\mu')\\ &+(1+\delta_{0n})^2 I_2^{n+}(z,\mu)\int_{-1}^{1}\mathrm{d}\mu'(\sigma_\mathrm{s}(z)P_{22}^n-\sigma_\mathrm{s}^0(z)P_{022}^n)I_2^n(z,\mu')\\ &+(1-\delta_{0n})I_3^{n+}(z,\mu)\int_{-1}^{1}\mathrm{d}\mu'(\sigma_\mathrm{s}(z)P_{31}^n-\sigma_\mathrm{s}^0(z)P_{031}^n)I_1^n(z,\mu')\\ &+(1-\delta_{0n})I_4^{n+}(z,\mu)\int_{-1}^{1}\mathrm{d}\mu'(\sigma_\mathrm{s}(z)P_{41}^n-\sigma_\mathrm{s}^0(z)P_{041}^n)I_1^n(z,\mu')\Big]\Big\},\end{aligned} \tag{35}$$

$$\begin{aligned}\Delta E_3 = -\pi\sum_{n=0}^{N}(1-\delta_{0n})\int_0^{z_\mathrm{T}}\mathrm{d}\mu\Big\{&(\sigma_\mathrm{t}(z)-\sigma_\mathrm{t}^0(z))I_3^{n+}(z,\mu)I_3^n(z,\mu)\\ &-\pi\Big[-I_1^{n+}(z,\mu)\int_{-1}^{1}\mathrm{d}\mu'(\sigma_\mathrm{s}(z)P_{13}^n-\sigma_\mathrm{s}^0(z)P_{013}^n)I_3^n(z,\mu')\end{aligned}$$

$$
\begin{aligned}
&-I_2^{n+}(z,\mu)\int_{-1}^{1} \mathrm{d}\mu'(\sigma_\mathrm{s}(z)P_{23}^n - \sigma_\mathrm{s}^0(z)P_{023}^n)I_3^n(z,\mu') \\
&+I_3^{n+}(z,\mu)\int_{-1}^{1} \mathrm{d}\mu'(\sigma_\mathrm{s}(z)P_{33}^n - \sigma_\mathrm{s}^0(z)P_{033}^n)I_3^n(z,\mu') \\
&+I_4^{n+}(z,\mu)\int_{-1}^{1} \mathrm{d}\mu'(\sigma_\mathrm{s}(z)P_{43}^n - \sigma_\mathrm{s}^0(z)P_{043}^n)I_3^n(z,\mu')\Bigg]\Bigg\}. \qquad (36)
\end{aligned}
$$

And finally

$$
\begin{aligned}
\Delta E_4 = &-\pi\sum_{n=0}^{N}(1-\delta_{0n})\int_0^{z_\mathrm{T}}\mathrm{d}z\int_{-1}^{1}\mathrm{d}\mu\Bigg\{(\sigma_\mathrm{t}(z)-\sigma_\mathrm{t}^0(z))I_4^{n+}(z,\mu)I_4^n(z,\mu) \\
&-\pi\Bigg[-I_1^{n+}(z,\mu)\int_{-1}^{1} \mathrm{d}\mu'(\sigma_\mathrm{s}(z)P_{14}^n - \sigma_\mathrm{s}^0(z)P_{014}^n)I_4^n(z,\mu') \\
&-I_2^{n+}(z,\mu)\int_{-1}^{1} \mathrm{d}\mu'(\sigma_\mathrm{s}(z)P_{24}^n - \sigma_\mathrm{s}^0(z)P_{024}^n)I_4^n(z,\mu') \\
&+I_3^{n+}(z,\mu)\int_{-1}^{1} \mathrm{d}\mu'(\sigma_\mathrm{s}(z)P_{34}^n - \sigma_\mathrm{s}^0(z)P_{034}^n)I_4^n(z,\mu') \\
&+I_4^{n+}(z,\mu)\int_{-1}^{1} \mathrm{d}\mu'(\sigma_\mathrm{s}(z)P_{44}^n - \sigma_\mathrm{s}^0(z)P_{044}^n)I_4^n(z,\mu')\Bigg]\Bigg\}. \qquad (37)
\end{aligned}
$$

To go any further requires assumptions as to the exact nature of the problem under investigation (for example, fluxes vs. radiances), and the form of the perturbation, which also relates to the mathematical structure of the scattering matrix.

5. Applications

The formalism just derived is completely general, and may therefore be applied to a wide variety of problems. Our ultimate aim is to apply this technique to the inversion of ground-based or satellite observations of both intensity and polarization, with the intention of extracting information about the optical properties of atmospheric aerosols, in an extension of the technique we developed for the inversion of radiance data [7,8]. An application such as this will require the full Fourier series for the measured quantities, as well as judicious choices as to the nature of the phase matrix, which has 16 components, compared to just one in the scalar case. For example, it will probably be more logical to consider specific aerosol models, with pre-computed phase matrices, rather than attempt to extract the expansion coefficients of the Legendre series [15].

A far simpler application is to azimuth independent quantities such as surface fluxes. Although there is usually very little difference between fluxes calculated using the vector versus the scalar radiative transfer equation, we feel that this is the logical place to start. In a follow-up paper, we apply this formalism to a series of realistic perturbations to an atmospheric model, for the computation of surface flux.

Appendix A

A.1. Proof of the adjoint property

We need to show that the adjoint transport operator, as defined, satisfies the definition of an adjoint, namely

$$\langle \mathbf{L}^{+}\vec{I}^{+},\vec{I}\rangle = \langle \vec{I}^{+},\mathbf{L}\vec{I}\rangle. \tag{A.1}$$

In writing this equation it is understood that an inner product is performed before the phase space integrals are performed. In order to simplify the analysis, we will examine this equation in pieces, by firstly proving that

$$E_1 \equiv \langle \mathbf{L}_1^{+}\vec{I}^{+},\vec{I}\rangle = \langle \vec{I}^{+},\mathbf{L}_1\vec{I}\rangle. \tag{A.2}$$

The transport operator is defined as

$$\mathbf{L} = \mu\frac{\partial}{\partial z}\mathbf{M} + \sigma_{\mathrm{t}}(z)\mathbf{M} - \sigma_{\mathrm{s}}(z)\int \mathrm{d}\Omega'\,\mathbf{P}(z,\Omega'\to\Omega)\circ \tag{A.3}$$

and its adjoint as

$$\mathbf{L}^{+} = -\mu\frac{\partial}{\partial z}\mathbf{M} + \sigma_{\mathrm{t}}(z)\mathbf{M} - \sigma_{\mathrm{s}}(z)\int \mathrm{d}\Omega'\,\tilde{\mathbf{P}}(z,\Omega'\to\Omega)\circ\,. \tag{A.4}$$

Note that we will find it convenient to use both a tilde, ∼, and a superscript T, to denote the transpose of a vector or matrix. To verify the above equation, we may therefore write

$$\mathbf{L}_1 = \begin{pmatrix} \mu\frac{\partial}{\partial z} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} + \mathbf{A}, \tag{A.5}$$

where $\mathbf{A}$ is a 4×4 matrix which does not contain any differential operators. (In fact, in this case it will contain only the first column.) Similarly we may write for the adjoint

$$\mathbf{L}_1^{+} = \begin{pmatrix} -\mu\frac{\partial}{\partial z} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} + \tilde{\mathbf{A}}. \tag{A.6}$$

Substituting these expressions we find that

$$E_1 = -\iint \mu \frac{\partial I_1^+}{\partial z} I_1 \, dz \, d\Omega + \langle \tilde{\mathbf{A}} \vec{I}^+, \vec{I} \rangle. \tag{A.7}$$

The first term in this expression may be integrated by parts to give

$$-\iint \mu \frac{\partial I_1^+}{\partial z} I_1 \, dz \, d\Omega = -\int \mu I_1^+ \, I_1 |_0^{z_T} \, d\Omega + \iint I_1^+ \mu \frac{\partial I_1}{\partial z} \, dz \, d\Omega. \tag{A.8}$$

Note that the first term is zero by virtue of the boundary conditions of no in-coming radiance quantities, and no out-going adjoint quantities. The remaining term in the expression for E_1 may be manipulated by standard matrix results as follows:

$$\langle \tilde{\mathbf{A}} \vec{I}^+, \vec{I} \rangle = \iint (\tilde{\mathbf{A}} \vec{I}^+)^{\mathrm{T}} \vec{I} \, dz \, d\Omega = \iint (\vec{I}^+)^{\mathrm{T}} \mathbf{A} \vec{I} \, dz \, d\Omega = \langle \vec{I}^+, \mathbf{A} \vec{I} \rangle. \tag{A.9}$$

Thus, we have shown that

$$E_1 = \iint I_1^+ \mu \frac{\partial I_1}{\partial z} \, dz \, d\Omega + \langle \vec{I}^+, \mathbf{A} \vec{I} \rangle = \langle \vec{I}^+, \mathbf{L}_1 \vec{I} \rangle. \tag{A.10}$$

Similar analysis applies to the other three components, so the result is proved in general.

A.2. Solving the adjoint equation

The polarized adjoint transfer equation may be written, in general, as

$$\mathbf{L}^+ \vec{I}^+(z, \Omega) = \vec{Q}^+(z, \Omega). \tag{A.11}$$

We now introduce a new vector

$$\vec{\Psi}(z, \Omega) = \vec{I}^+(z, -\Omega) \tag{A.12}$$

and substitute it into the adjoint transport equation:

$$\begin{aligned} -\mu \frac{\partial \vec{\Psi}(z, -\Omega)}{\partial z} &+ \sigma_t(z) \vec{\Psi}(z, -\Omega) - \sigma_s(z) \int_{4\pi} d\Omega' \, \tilde{\mathbf{P}}(z, \Omega \to \Omega') \vec{\Psi}(z, -\Omega') \\ &= \vec{Q}^+(z, \Omega). \end{aligned} \tag{A.13}$$

If we now change the direction of Ω to $-\Omega$, we obtain

$$\begin{aligned} \mu \frac{\partial \vec{\Psi}(z, \Omega)}{\partial z} &+ \sigma_t(z) \vec{\Psi}(z, \Omega) - \sigma_s(z) \int_{4\pi} d\Omega' \, \tilde{\mathbf{P}}(z, -\Omega \to \Omega') \vec{\Psi}(z, -\Omega') \\ &= \vec{Q}^+(z, -\Omega). \end{aligned} \tag{A.14}$$

We now also change the direction of Ω' to $-\Omega'$, so that

$$\begin{aligned} \mu \frac{\partial \vec{\Psi}(z, \Omega)}{\partial z} &+ \sigma_t(z) \vec{\Psi}(z, \Omega) - \sigma_s(z) \int_{4\pi} d\Omega' \, \tilde{\mathbf{P}}(z, -\Omega \to -\Omega') \vec{\Psi}(z, \Omega') \\ &= \vec{Q}^+(z, -\Omega). \end{aligned} \tag{A.15}$$

Finally, we use the reciprocity relation for scattering, namely

$$\tilde{\mathbf{P}}(z,-\Omega \to -\Omega') = \mathbf{P}(z,\Omega' \to \Omega) \tag{A.16}$$

to obtain the result

$$\mu\frac{\partial\vec{\Psi}(z,\Omega)}{\partial z} + \sigma_{\mathrm{s}}(z)\vec{\Psi}(z,\Omega) - \sigma_{\mathrm{t}}(z)\int_{4\pi}\mathrm{d}\Omega'\,\mathbf{P}(z,\Omega' \to \Omega)\vec{\Psi}(z,\Omega') = \vec{Q}^{+}(z,-\Omega). \tag{A.17}$$

Thus we see that, exactly as in the scalar case [11], we may obtain the adjoint (vector) field by solving the standard (or 'forward') transfer equation for an angle reversed source to obtain the (vector) pseudo fields, and then reversing directions to obtain the final adjoint fields.

A.3. Derivation of the perturbation equation

Derivation of the standard perturbation equation follows essentially the same steps as for the scalar case [1]. We start by writing

$$\begin{aligned}\vec{Q} &= \mathbf{L}\vec{I} = (\mathbf{L}_0 + \Delta\mathbf{L})(\vec{I}_0 + \Delta\vec{I})\\ &= \mathbf{L}_0\vec{I}_0 + \Delta\mathbf{L}\vec{I}_0 + \mathbf{L}\Delta\vec{I}.\end{aligned} \tag{A.18}$$

Since we also know that

$$\vec{Q} = \mathbf{L}_0\vec{I}_0. \tag{A.19}$$

We obtain by subtracting these equations

$$\Delta\mathbf{L}\vec{I}_0 + \mathbf{L}\Delta\vec{I} = 0. \tag{A.20}$$

This equation must be true in quite general circumstances, and so it must also be valid when expressed in terms of a single component of the transport operator, that is

$$\Delta\mathbf{L}_i\vec{I}_0 + \mathbf{L}_i\Delta\vec{I} = 0. \tag{A.21}$$

We now multiply this equation by the base case adjoint (vector) field, and integrate, using the definitions introduced above, to obtain

$$\langle\vec{I}_0^{+}, \Delta\mathbf{L}_i\vec{I}_0\rangle + \langle\vec{I}_0^{+}, \mathbf{L}_i\Delta\vec{I}\rangle = 0. \tag{A.22}$$

This equation may now be manipulated as follows:

$$\begin{aligned}0 &= \langle\vec{I}_0^{+}, \Delta\mathbf{L}_i\vec{I}_0\rangle + \langle\vec{I}^{+} - \Delta\vec{I}^{+}, \mathbf{L}_i\Delta\vec{I}\rangle\\ &= \langle\vec{I}_0^{+}, \Delta\mathbf{L}_i\vec{I}_0\rangle + \langle\vec{I}^{+}, \mathbf{L}_i(\vec{I} - \vec{I}_0)\rangle - \langle\Delta\vec{I}^{+}, \mathbf{L}_i\Delta\vec{I}\rangle \qquad \text{(A.23)}\\ &= \langle\vec{I}_0^{+}, \Delta\mathbf{L}_i\vec{I}_0\rangle + \langle\vec{I}^{+}, \mathbf{L}_i\vec{I}\rangle - \langle\vec{I}^{+}, \mathbf{L}_i\vec{I}_0\rangle - \langle\Delta\vec{I}^{+}, \mathbf{L}_i\Delta\vec{I}\rangle. \qquad \text{(A.24)}\end{aligned}$$

Now we know that

$$E_i = \langle\vec{I}^{+}, \mathbf{L}_i\vec{I}\rangle \tag{A.25}$$

and also that

$$E_{0i} = \langle \vec{R}, \vec{I}_0 \rangle = \langle \mathbf{L}_i^+ \vec{I}^+, \vec{I}_0 \rangle = \langle \vec{I}^+, \mathbf{L}_i \vec{I}_0 \rangle. \tag{A.26}$$

Hence, we may combine these results to obtain

$$E_i = E_{0i} - \langle \vec{I}_0^+, \Delta \mathbf{L}_i \vec{I}_0 \rangle + \langle \Delta \vec{I}^+, \mathbf{L}_{0i} \Delta \vec{I} \rangle. \tag{A.27}$$

Finally, in the spirit of (first order) perturbation theory, we assume that we may neglect a term which contains the product of two small factors (“second order of smallness”), to finally obtain

$$E_i \cong E_{0i} - \langle \vec{I}_0^+, \Delta \mathbf{L} \vec{I}_0 \rangle. \tag{A.28}$$

References

[1] Box MA, Gerstl SAW, Simmer C. Beitr Phys Atmos 1989;62:193–9.
[2] Box MA, Croke B, Gerstl SAW, Simmer C. Beitr Phys Atmos 1989;62:200–11.
[3] Trautmann T, Box MA. J Geophys Res 1995;100:1081–92.
[4] Box MA, Beck S, Trautmann T. J Geophys Res 1996;101:19,293–7.
[5] Box MA, Loughlin PE, Samaras M, Trautmann T. J Geophys Res 1997;102:4333–42.
[6] Loughlin PE, Box MA. Photochem Photobiol B 1998;43:73–85.
[7] Box MA, Sendra C. Appl Opt 1999;38:1636–43.
[8] Sendra C, Box MA. JQSRT 2000;64:499–515.
[9] Holben BN, Eck TF, Slutsker I, Tanre D, Buis JP, Setzer A, Vermote E, Reagan JA, Kaufman YJ, Nakajima T, Lavenu F, Jankowiak I, Smirnov A. Remote Sens Environ 1998;66:1–16.
[10] Herman M, Deuze JL, Devaux C, Goloub P, Breon FM, Tanre D. J Geophys Res 1997;102:17,039–49.
[11] Box MA, Gerstl SAW, Simmer C. Beitr Phys Atmos 1988;61:303–11.
[12] Chandrasekhar S. Radiative transfer. New York: Dover, 1960.
[13] Coulson KL. Polarization and intensity of light in the atmosphere. Hampton, VA: A. Deepak Publishing, 1988.
[14] Dave JV. Appl Opt 1970;9:2673–84.
[15] Polonsky I, Box MA. J Atmos Sci, accepted for publication.