5 Linearization of vector radiative transfer by means of the forward-adjoint perturbation theory and its use in atmospheric remote sensing

Otto P. Hasekamp and Jochen Landgraf

5.1 Introduction

Aerosols directly affect the Earth's climate by scattering and absorption of radiation, and indirectly by changing the microphysical properties of clouds. The total effect of aerosols on climate is very uncertain, both in magnitude and even in sign, representing one of the largest uncertainties in climate research. In order to improve our understanding of the effect of aerosols on climate, global measurements are needed of a number of aerosol properties such as size of particles, their refractive index and aerosol optical thickness. The only way to obtain these parameters at a global scale is by means of satellite remote sensing.

Information on aerosol properties is contained in the spectral and angular behavior of the total intensity and the polarization properties of backscattered sunlight. Most satellite instruments that are used for aerosol retrieval only measure the intensity spectrum of backscattered light. Among these instruments are the Advanced Very High Resolution Radiometer (AVHRR), the Moderate Resolution Imaging Spectroradiometer (MODIS), the Total Ozone Mapping Spectrometer (TOMS), the Global Ozone Monitoring Experiment (GOME), the Scanning Imaging Absorption Spectrometer for Atmospheric Chartography (SCIAMACHY), the Medium Resolution Imaging Sensor (MERIS), and the Ozone Monitoring Instrument (OMI). Retrieval algorithms for instruments of this type allow the choice between a number of standard aerosol models (a combination of size distribution and refractive index), where the model that agrees best with the measured spectrum is used to determine the aerosol optical thickness. On the one hand these intensity -only retrievals do not provide enough information to answer the relevant questions in climate research, and on the other hand the retrieved optical thickness depends critically on the choice of the aerosol model. The information content with respect to aerosol properties is significantly larger for multiple-viewing-angle intensity measurements as performed by the Multi-angle Imaging Spectro-Radiometer (MISR), for multiple-viewing-angle intensity and polarization measurements as performed by the Polarization and Anisotropy of Reflectances for Atmospheric Sciences Coupled with Observations from a Lidar instrument (PARASOL), and for single-viewing-angle intensity and polarization measurements, as performed by GOME-2. These instruments contain enough information to do more than just distinguishing between a number of aerosol models. Therefore, the 'classical' retrieval approach described above is not sufficient for these more advanced instruments. Instead, a retrieval approach is required that makes full use of the information content of the measurement.

Recently a new approach to the retrieval of aerosol properties has been developed (Hasekamp and Landgraf, 2005a,b). Instead of assuming a number of standard aerosol models, the developed method aims to retrieve microphysical aerosol properties corresponding to a bi-modal aerosol size distribution. The retrieval of these aerosol parameters from satellite measurements requires a forward model $\mathbf{F}$ that describes how the measured data depend on the aerosol parameters, viz.

$$\mathbf{y} = \mathbf{F}(\mathbf{x}) + \mathbf{e}_y. \tag{5.1}$$

Here $\mathbf{y}$ is the measurement vector containing the measured data, e.g. intensity and/or polarization measurements at different wavelengths and/or different viewing angles, and $\mathbf{e}_y$ is the corresponding error vector. $\mathbf{x}$ is the state vector containing the aerosol parameters to be retrieved. The forward model consists of two parts. The first part relates the physical aerosol properties (size distribution, refractive index) to their optical properties (scattering and extinction coefficients, phase matrix). This relation can be described by Mie theory for spherical particles (van der Hulst, 1957) or alternative theories for particles of other shapes (Wiscombe and Grams, 1988; Koepke and Hess, 1988; Mishchenko and Travis, 1994; Mishchenko et al., 1995). The second part of the forward model is an atmospheric radiative transfer model that simulates the intensity vector at the top of the atmosphere for given optical input parameters. Since the forward model $\mathbf{F}$ is non-linear in the microphysical aerosol parameters contained in $\mathbf{x}$, the inversion of Eq. (5.1) has to be performed iteratively. Hereto, the forward model $\mathbf{F}$ in Eq. (5.1) is replaced by its linear approximation in each iteration step n

$$\mathbf{y} \approx \mathbf{F}(\mathbf{x}_\mathrm{n}) + \mathbf{K}\ (\mathbf{x} - \mathbf{x}_n) + \mathbf{e}_y, \tag{5.2}$$

where $\mathbf{x}_\mathrm{n}$ is the state vector for the iteration step under consideration and $\mathbf{K}$ is the Jacobian matrix containing the derivatives of the forward model with respect to the elements of $\mathbf{x}$, where element K_{ij} of $\mathbf{K}$ is defined by:

$$K_{ij} = \frac{\partial F_i}{\partial x_j}(\mathbf{x}_\mathrm{n}). \tag{5.3}$$

The inversion of Eq. (5.2) can be performed analytically. Once $\mathbf{x}_n$ is close enough to the true state vector $\mathbf{x}$, the Jacobian matrix $\mathbf{K}$ can be used to calculate the mapping of the measurement errors $\mathbf{e}_y$ to errors on the retrieved aerosol parameters $\mathbf{e}_x$. Thus, the Jacobian matrix $\mathbf{K}$ plays an important role in the retrieval process, both for finding an appropriate solution of the inversion problem and for a solid error analysis. Therefore, in the most general case of aerosol retrieval from intensity and polarization measurements, a linearized vector radiative transfer model is needed that simulates the intensity vector at the top of the model atmosphere and additionally calculates the derivatives of the Stokes parameters with respect to the aerosol properties to be retrieved.

For scalar radiative transfer, a general linearization approach was proposed by Marchuk (1964) who employed the forward-adjoint perturbation theory approach, known from neutron transport theory (see, for example, Bell and Classtone (1970)), to atmospheric scalar radiative transfer. This approach has been used by, for example, Ustinov (1991), Rozanov et al. (1998) and Landgraf et al. (2002) for the linearization of scalar radiative transfer with respect to atmospheric absorption properties, by Landgraf et al. (2001) for the linearization with respect to surface properties, and by, for example, Ustinov (1992) and Sendra and Box (2000) for the linearization with respect to atmospheric scattering properties. Another approach for the linearization of scalar radiative transfer has been followed by Spurr et al. (2001), who developed an analytical linearization with respect to absorption and scattering properties for the discrete ordinate method of scalar radiative transfer (Chandrasekhar, 1960; Stamnes et al., 1988). For plane-parallel vector radiative transfer, an analytical linearization with respect to atmospheric absorption properties has been developed by Hasekamp and Landgraf (2002), who extended the forward-adjoint perturbation theory to include polarization.

An analytical linearization of vector radiative transfer with respect to atmospheric scattering properties has recently been achieved by Hasekamp and Landgraf (2005a). They combined the linearization of radiative transfer with a linearization of Mie theory to obtain a radiative transfer model that provides the requested derivatives of the Stokes parameters at the top of the atmosphere with respect to microphysical aerosol properties. Based on this linearized vector radiative transfer model, Hasekamp and Landgraf (2005b) developed a novel approach to the retrieval of microphysical aerosol properties from multi-spectral single-viewing-angle measurements of intensity and polarization. This retrieval approach uses the Phillips–Tikhonov regularization method for the analytical inversion of the linearized radiative transfer model. A powerful feature of this approach is that it quantifies the information content of the measurement as part of the retrieval procedure and extracts the available information. The aim of this chapter is to review this retrieval approach based on linearized radiative transfer and Phillips–Tikhonov regularization. The first part of this chapter (section 5.2–5.5) is devoted to the linearization of vector radiative transfer with respect to microphysical aerosol properties. The second part of the chapter (section 5.6 and 5.7) discusses the application of the linearized vector radiative transfer model in a retrieval scheme using Phillips–Tikhonov regularization.

5.2 Radiative transfer model

The radiance and state of polarization of light at a given wavelength can be described by an intensity vector **I** which has the Stokes parameters as its components (see, for example, Chandrasekhar (1960)):

$$\mathbf{I} = [I, Q, U, V]^T \,, \tag{5.4}$$

where T indicates the transposed vector, and the Stokes parameters are defined with respect to a certain reference plane. The angular dependence of single scattering of polarized light can be described by means of the scattering phase matrix $\mathbf{P}$. We will restrict ourselves to scattering phase matrices of the form

$$\mathbf{P}(\theta) = \begin{pmatrix} p_1(\theta) & p_5(\theta) & 0 & 0 \\ p_5(\theta) & p_2(\theta) & 0 & 0 \\ 0 & 0 & p_3(\theta) & p_6(\theta) \\ 0 & 0 & -p_6(\theta) & p_4(\theta) \end{pmatrix} . \tag{5.5}$$

where $p_1, p_2, \ldots, p_6$ are certain functions of scattering angle θ and the scattering plane is the plane of reference. This type of scattering matrix is valid for (see, for example, van de Hulst (1957)) (i) scattering by an assembly of randomly oriented particles each having a plane of symmetry, (ii) scattering by an assembly containing particles and their mirror particles in equal numbers and with random orientations, (iii) Rayleigh scattering with or without depolarization effects.

To discuss the single scattering properties of aerosol particles we will use the scattering plane as the plane of reference. However, for the atmospheric radiative transfer calculations in this chapter we will use the local meridian plane, defined as the plane going through the direction of propagation and the vertical direction, as reference plane.

5.2.1 Radiative transfer equation in operator form

We consider a plane-parallel, macroscopically isotropic atmosphere bounded below by a reflecting surface. Furthermore, we ignore inelastic scattering and thermal emission. The equation of transfer for polarized light is now given in its forward formulation by

$$\hat{\mathbf{L}}\,\mathbf{I} = \mathbf{S}, \tag{5.6}$$

where the transport operator

$$\begin{aligned} \hat{\mathbf{L}} \;=\; & \int_{4\pi} \mathrm{d}\tilde{\Omega} \Bigg\{ \left[\mu \frac{\partial}{\partial z} + K_{\text{ext}}(z) \right] \delta(\mathbf{\Omega} - \tilde{\mathbf{\Omega}})\, \mathbf{E} - \\ & \frac{K_{\text{sca}}(z)}{4\pi}\, \mathbf{Z}(z, \tilde{\mathbf{\Omega}}, \mathbf{\Omega}) - \delta(z)\, \Theta(\mu)|\mu|\, \mathbf{R}_s(\tilde{\mathbf{\Omega}}, \mathbf{\Omega})\, \Theta(-\tilde{\mu})|\tilde{\mu}| \Bigg\}, \end{aligned} \tag{5.7}$$

is adopted from scalar radiative transfer (Marchuk, 1964; Box et al., 1988; Ustinov, 2001; Landgraf et al., 2002). Here, z describes altitude, the direction $\mathbf{\Omega}$ is given by (μ, φ) where φ is the azimuthal angle measured clockwise when looking downward and μ is the cosine of the zenith angle ($\mu < 0$ for downward directions and $\mu > 0$ for upward directions). Furthermore, $d\Omega = d\mu\; d\varphi$, $\mathbf{E}$ is the 4×4 unity matrix, K_{ext} and K_{sca} represent the extinction and scattering coefficients, respectively, Θ represents the Heaviside step function, and δ is the Dirac-delta function with $\delta(\mathbf{\Omega} - \tilde{\mathbf{\Omega}}) = \delta(\mu - \tilde{\mu})\delta(\varphi - \tilde{\varphi})$. The first term of the radiative transfer operator describes the extinction of light, whereas the second term represents scattering of light from direction $\tilde{\mathbf{\Omega}}$ to $\mathbf{\Omega}$ with the phase matrix $\mathbf{Z}(z,\tilde{\mathbf{\Omega}},\mathbf{\Omega})$, defined with respect to the local meridian plane. The last term on the right-hand side of Eq. (5.7) describes the surface reflection at the lower boundary of the atmosphere with reflection matrix $\mathbf{R}_s$.

The right-hand side of Eq. (5.6) provides the source of light and can either be a volume source inside the atmosphere or a surface source chosen to reproduce the incident flux conditions at the boundaries of the atmosphere, or some combination of the two. In the UV and visible part of the spectrum the radiation source $\mathbf{S}$ is determined by the unpolarized sunlight that illuminates the top of the Earth atmosphere:

$$\mathbf{S}(z,\mathbf{\Omega}) = \mu_o \delta(z - z_{\text{top}})\delta(\mathbf{\Omega} - \mathbf{\Omega}_o)\mathbf{F}_o. \tag{5.8}$$

Here, z_{top} is the height of the model atmosphere, $\mathbf{\Omega}_o = (-\mu_o, \varphi_o)$ describes the geometry of the incoming solar beam (we define $\mu_o > 0$), and $\mathbf{F}_o$ is given by

$$\mathbf{F}_o = [F_o, 0, 0, 0]^T , \tag{5.9}$$

where F_o is the solar flux per unit area perpendicular to the direction of the solar beam. Because the reflection of light at the ground surface is already included in the radiative transfer operator (5.7) and the incoming solar beam is represented by the radiation source of Eq. (5.8), the intensity vector $\mathbf{I}$ is subject to homogeneous boundary conditions:

$$\begin{aligned} \mathbf{I}(z_{\text{top}},\mathbf{\Omega}) &= [0,0,0,0]^T \qquad \text{for } \mu < 0, \\ \mathbf{I}(0,\mathbf{\Omega}) &= [0,0,0,0]^T \qquad \text{for } \mu > 0. \end{aligned} \tag{5.10}$$

In conjunction with these boundary conditions, the radiation source $\mathbf{S}$ can be interpreted as located at a vanishingly small distance below the upper boundary. Similarly, the surface reflection takes place at a vanishingly small distance above the lower boundary (see, for example, Morse and Feshbach (1953)).

In order to handle the integration over azimuth angle in Eq. (5.6) we use a decomposition of the radiative transfer equation into corresponding equations per Fourier component (Hovenier and van der Mee, 1983; de Haan et al., 1987):

$$\hat{\mathbf{L}}^m \, \mathbf{I}^{\pm m} = \mathbf{S}^{\pm m}, \tag{5.11}$$

with

$$\begin{aligned} \hat{\mathbf{L}}^m \;=\; & \int_{-1}^{1} d\tilde{\mu} \Bigg\{ \left[\mu \frac{\partial}{\partial z} + K_{\text{ext}}(z) \right] \delta(\mu - \tilde{\mu}) \, \mathbf{E} - \\ & \frac{K_{\text{sca}}(z)}{2} \, \mathbf{Z}^m(z,\tilde{\mu},\mu) - \delta(z) \, \Theta(\mu)|\mu| \, \mathbf{R}_s^m(\tilde{\mu},\mu) \, \Theta(-\tilde{\mu})|\tilde{\mu}| \Bigg\}. \end{aligned} \tag{5.12}$$

The corresponding Fourier expansion of the intensity vector is given by

$$\mathbf{I}(z,\mathbf{\Omega}) = \sum_{m=0}^{\infty}(2-\delta_{m0})\left[\mathbf{B}^{+m}(\varphi_o-\varphi)\mathbf{I}^{+m}(z,\mu)+\mathbf{B}^{-m}(\varphi_o-\varphi)\mathbf{I}^{-m}(z,\mu)\right], \tag{5.13}$$

where δ_{m0} is the Kronecker delta, and

$$\mathbf{B}^{+m}(\varphi) = \mathrm{diag}[\cos m\varphi, \cos m\varphi, \sin m\varphi, \sin m\varphi], \tag{5.14}$$

$$\mathbf{B}^{-m}(\varphi) = \mathrm{diag}[-\sin m\varphi, -\sin m\varphi, \cos m\varphi, \cos m\varphi]. \tag{5.15}$$

The Fourier coefficients of the intensity vector are given by

$$\begin{aligned}\mathbf{I}^{+m}(z,\mu) &= \frac{1}{2\pi}\int_0^{2\pi} d\varphi\, \mathbf{B}^{+m}(\varphi_o-\varphi)\, \mathbf{I}(z,\mathbf{\Omega}),\\ \mathbf{I}^{-m}(z,\mu) &= \frac{1}{2\pi}\int_0^{2\pi} d\varphi\, \mathbf{B}^{-m}(\varphi_o-\varphi)\, \mathbf{I}(z,\mathbf{\Omega}).\end{aligned} \tag{5.16}$$

Similarly, a Fourier expansion of the radiation source $\mathbf{S}(z,\mathbf{\Omega})$ is obtained, with Fourier coefficients

$$\begin{aligned}\mathbf{S}^{+m}(z,\mu) &= \frac{1}{2\pi}\,\mu_o\,\delta(z-z_{top})\,\delta(\mu_o+\mu)\,\mathbf{F}_o,\\ \mathbf{S}^{-m}(z,\mu) &= [0,0,0,0]^T.\end{aligned} \tag{5.17}$$

From Eqs (5.10) and (5.17) it follows that $\mathbf{I}^{-m}=0$, so the Fourier expansion of the intensity vector contains terms of $\mathbf{I}^{+m}$ only.

The Fourier expansion of the phase matrix is given by

$$\begin{aligned}\mathbf{Z}(z,\tilde{\mathbf{\Omega}},\mathbf{\Omega}) = \frac{1}{2}\sum_{m=0}^{\infty}(2-\delta_{m0})\big[&\mathbf{B}^{+m}(\tilde{\varphi}-\varphi)\mathbf{Z}^{m}(z,\tilde{\mu},\mu)(\mathbf{E}+\mathbf{\Lambda})\\ &+\mathbf{B}^{-m}(\tilde{\varphi}-\varphi)\mathbf{Z}^{m}(z,\tilde{\mu},\mu)(\mathbf{E}-\mathbf{\Lambda})\big],\end{aligned} \tag{5.18}$$

where

$$\mathbf{\Lambda} = \mathrm{diag}\,[1,1,-1,-1]\,. \tag{5.19}$$

The mth Fourier coefficient of the phase matrix can be calculated by

$$\mathbf{Z}^{m}(z,\tilde{\mu},\mu) = (-1)^m \sum_{l=m}^{L} \mathbf{P}^l_m(-\mu)\,\mathbf{S}^l(z)\,\mathbf{P}^l_m(-\tilde{\mu}), \tag{5.20}$$

where L is a suitable truncation index (Ustinov, 1988) and $\mathbf{P}^l_m$ is the generalized spherical function matrix given by

$$\mathbf{P}^m_l(\mu) = \begin{pmatrix} P^l_{m0}(\mu) & 0 & 0 & 0\\ 0 & P^l_{m+}(\mu) & P^l_{m-}(\mu) & 0\\ 0 & P^l_{m-}(\mu) & P^l_{m+}(\mu) & 0\\ 0 & 0 & 0 & P^l_{m0}(\mu)\end{pmatrix}, \tag{5.21}$$

where

$$P^l_{m\pm} = \frac{1}{2}\left(P^l_{m,-2} \pm P^l_{m,2}\right), \tag{5.22}$$

and $P^l_{mn}(\mu)$ are the generalized spherical functions (Gel'fand et al., 1963), which were introduced in atmospheric radiative transfer by Kuščer and Ribarič (1959). $\mathbf{S}^l$ is the expansion coefficient matrix having the form

$$\mathbf{S}^l = \begin{pmatrix} \alpha^l_1 & \alpha^l_5 & 0 & 0 \\ \alpha^l_5 & \alpha^l_2 & 0 & 0 \\ 0 & 0 & \alpha^l_3 & \alpha^l_6 \\ 0 & 0 & -\alpha^l_6 & \alpha^l_4 \end{pmatrix}, \tag{5.23}$$

where the expansion coefficients follow from the scattering phase matrix $\mathbf{P}$ in Eq. (5.5) (see, for example, de Rooij and van der Stap (1984)):

$$\alpha^l_1 = \frac{2l+1}{2} \int_{-1}^{1} P^l_{0,0}(\cos\theta) p_1(\theta)\, d(\cos\theta), \tag{5.24}$$

$$\alpha^l_2 + \alpha^l_3 = -\frac{2l+1}{2}\sqrt{\frac{(l-2)!}{(l+2)!}} \int_{-1}^{1} P^l_{2,2}(\cos\theta)(p_2(\theta)+p_3(\theta))\, \mathrm{d}(\cos\theta), \tag{5.25}$$

$$\alpha^l_2 - \alpha^l_3 = -\frac{2l+1}{2}\sqrt{\frac{(l-2)!}{(l+2)!}} \int_{-1}^{1} P^l_{2,-2}(\cos\theta)(p_2(\theta)+p_3(\theta))\, \mathrm{d}(\cos\theta), \tag{5.26}$$

$$\alpha^l_4 = \frac{2l+1}{2} \int_{-1}^{1} P^l_{0,0}(\cos\theta) p_4(\theta)\, \mathrm{d}(\cos\theta), \tag{5.27}$$

$$\alpha^l_5 = \frac{2l+1}{2} \int_{-1}^{1} P^l_{0,2}(\cos\theta) p_5(\theta)\, \mathrm{d}(\cos\theta), \tag{5.28}$$

$$\alpha^l_6 = \frac{2l+1}{2} \int_{-1}^{1} P^l_{0,2}(\cos\theta) p_6(\theta)\, \mathrm{d}(\cos\theta). \tag{5.29}$$

In this chapter we assume that the surface reflection matrix obeys the same symmetry relations as the scattering phase matrix (Hovenier, 1969) and thus can also be expanded in a Fourier series. The Fourier coefficients $\mathbf{R}^m_s$ of the surface reflection matrix $\mathbf{R}_s$ are given by:

$$\mathbf{R}^m_s(\tilde{\mu},\mu) = \frac{1}{2\pi}\int_0^{2\pi} \mathrm{d}(\tilde{\varphi}-\varphi)\left[\mathbf{B}^{+m}(\tilde{\varphi}-\varphi) + \mathbf{B}^{-m}(\tilde{\varphi}-\varphi)\right]\, \mathbf{R}_s(\tilde{\mathbf{\Omega}},\mathbf{\Omega}). \tag{5.30}$$

In order to obtain the Fourier coefficients $\mathbf{I}^m$ from Eq. (5.11) for a vertically inhomogeneous atmosphere, the model atmosphere has to be divided in a number of homogeneous layers, where each layer is characterized by a height-independent scattering coefficient, extinction coefficient, and scattering matrix. Several numerical models exist to solve the corresponding radiative transfer problem. We will use the Gauss–Seidel model described by Landgraf et al. (2002) and Hasekamp and Landgraf (2002).

5.3 Mie scattering calculations

The optical input parameters of the radiative transfer equation (5.6) are the extinction and scattering coefficients and the phase matrix $\mathbf{Z}$ in the form of expansion coefficient matrices $\mathbf{S}^l$. These parameters are determined by scattering and absorption by aerosols, scattering by air molecules, and absorption by atmospheric gases, and are obtained from these different components:

$$K_{\text{ext}} = K^a_{\text{ext}} + K^r_{\text{ext}} + K^g_{\text{ext}}, \tag{5.31}$$

$$K_{\text{sca}} = K^a_{\text{sca}} + K^r_{\text{sca}}, \tag{5.32}$$

$$\alpha = \frac{K^a_{\text{sca}}\,\alpha^a}{K_{\text{sca}}} + \frac{K^r_{\text{sca}}\,\alpha^r}{K_{\text{sca}}}, \tag{5.33}$$

where the superscript a denotes aerosol, the superscript r denotes Rayleigh scattering, the superscript g denotes gas absorption, and we omitted the sub- and superscripts of the expansion coefficients α^l_j.

The optical properties of aerosols depend on the size, shape, and type of aerosols. In this chapter we restrict ourselves to spherical aerosols which means that the optical properties of aerosols can be calculated using Mie theory. Here we will summarize the most important formulas needed for Mie calculations (see, for example, de Rooij and van der Stap (1984)). A complete overview of Mie scattering theory is given by van de Hulst (1957).

In order to calculate the elements of the Mie scattering phase matrix $\mathbf{P}$ in Eq. (5.5), we first consider the transformation matrix $\boldsymbol{\mathcal{F}}$ (van de Hulst, 1957) which is defined as

$$\boldsymbol{\mathcal{F}} = \frac{k^2\,C_{\text{sca}}}{4\pi}\,\mathbf{P}, \tag{5.34}$$

with elements $f_1, f_2, \ldots, f_6$, analogous to the elements $p_1, p_2, \ldots, p_6$ in Eq. (5.5). In Eq. (5.34) C_{sca} is the scattering cross-section, and $k = 2\pi/\lambda$, where λ denotes wavelength. For a single sphere of radius r the elements of the transformation matrix $\boldsymbol{\mathcal{F}}$ are given by

$$f_1 = \frac{1}{2}\,(S_1 S_1^* + S_2 S_2^*)\,, \tag{5.35}$$

$$f_2 = f_1, \tag{5.36}$$

$$f_3 = \frac{1}{2}\,(S_1 S_2^* + S_2 S_1^*)\,, \tag{5.37}$$

$$f_4 = f_3, \tag{5.38}$$

$$f_5 = \frac{1}{2}\,(S_1 S_1^* - S_2 S_2^*)\,, \tag{5.39}$$

$$f_6 = \frac{i}{2}\,(S_1 S_2^* - S_2 S_1^*)\,, \tag{5.40}$$

where we omitted the dependence on scattering angle θ and particle radius r. In Eqs (5.35)–(5.40), S_1 and S_2 are the elements of the two-by-two scattering amplitude matrix relating the electric field vector (containing the component parallel and the component perpendicular to the scattering plane) of the scattered beam to that of the incoming beam, and the asterisk denotes the complex conjugate. The functions S_1 and S_2 are given by

$$S_1(\theta) = \sum_{n=1}^{\infty} \frac{2n+1}{n(n+1)} \left(a_n \pi_n(\theta) + b_n \tau_n(\theta) \right), \tag{5.41}$$

$$S_2(\theta) = \sum_{n=1}^{\infty} \frac{2n+1}{n(n+1)} \left(b_n \pi_n(\theta) + a_n \tau_n(\theta) \right), \tag{5.42}$$

where π_n and τ_n are functions of only scattering angle and are expressed in associated Legendre functions as

$$\pi_n(\theta) = \frac{1}{\sin\theta} P_n^1(\cos\theta), \tag{5.43}$$

$$\tau_n(\theta) = \frac{d}{d\theta} P_n^1(\cos\theta). \tag{5.44}$$

The most substantial part of the Mie calculations is the computation of the Mie coefficients a_n and b_n in Eqs (5.41) and (5.42) which are functions of the particle's complex refractive index $m = m_r + im_i$ and the size parameter kr. A numerical procedure for calculating a_n and b_n is given by Rooij and van der Stap (1984), and is summarized in Appendix A of this chapter. The scattering and extinction cross-sections, $C_{\rm sca}$ and $C_{\rm ext}$, of a single sphere with radius r can also be calculated using the coefficients a_n and b_n:

$$C_{\rm sca}(r) = \frac{2\pi}{k^2} \sum_{n=1}^{\infty} (2n+1)[|a_n|^2 + |b_n|^2], \tag{5.45}$$

$$C_{\rm ext}(r) = \frac{2\pi}{k^2} \sum_{n=1}^{\infty} (2n+1) \mathrm{Re}(a_n + b_n). \tag{5.46}$$

Equations (5.35)–(5.42), (5.45), and (5.46) provide expressions for the scattering matrix and absorption and extinction cross-sections for a single sphere. In nature, a distribution of particles with different sizes is normally encountered. Under the assumption of independent scattering (see, for example, Hansen and Travis (1974)) the scattering and extinction cross-section for a size distribution are given by

$$\bar{C}_{\rm sca} = \int_0^{\infty} C_{\rm sca}(r)\; n(r)\; dr, \tag{5.47}$$

$$\bar{C}_{\rm ext} = \int_0^{\infty} C_{\rm ext}(r)\; n(r)\; dr, \tag{5.48}$$

where $n(r)$ is the aerosol size distribution normalized to unity (e.g. a lognormal distribution; see Appendix B).

Similarly, element f_j of the transformation matrix $\boldsymbol{\mathcal{F}}$ (5.34) for a size distribution is given by

$$\bar{f}_j = \int_0^\infty f_j(r)\, n(r)\,\mathrm{d}r. \tag{5.49}$$

In this chapter the integrations over size distribution are approximated by a sum over different size bins. Here, for a function $g(r)$ the integral over size distribution is approximated by

$$\bar{g} \approx \sum_{i=1}^{N} g(r_i)\, n_i\, \Delta r_i, \tag{5.50}$$

where r_i is the middle of size interval i, $n_i = n(r_i)$ and Δr_i is the width of size interval i.

The elements p_j of the scattering phase matrix $\mathbf{P}$ of an ensemble with given size distribution can be obtained from $\bar{f}_j$ and $\bar{C}_{\mathrm{sca}}$ via

$$p_j = \frac{4\pi}{k^2} \frac{\bar{f}_j}{\bar{C}_{\mathrm{sca}}}. \tag{5.51}$$

The expansion coefficients α_i^j can now be calculated from the elements of the scattering phase matrix $\mathbf{P}$ using Eqs (5.24)–(5.29). Here, the integrals over $\cos\theta$ can be calculated analytically (Domke, 1975) or numerically (de Rooij and van der Stap, 1984). The latter approach will be adopted for the calculations in this chapter. Furthermore, the aerosol extinction and scattering coefficients K^a_{ext} and K^a_{sca} for each homogeneous layer are obtained by multiplying the corresponding cross-section by the layer integrated aerosol number concentration in that layer.

5.4 Linearization of the forward model

As follows from the previous section, the relevant optical properties of spherical homogeneous aerosol particles can be obtained from the aerosol size distribution, the aerosol number concentration, and the aerosol refractive index. Often, the size distribution consists of different modes, where each mode contains particles of the same refractive index. Thus, in the most general case the elements of the state vector $\mathbf{x}$ for aerosol retrieval in Eq. (5.1) are for each homogeneous layer and mode of the size distribution, the real part of the refractive index m_r, the imaginary part of the refractive index m_i, the elements n_i of the discretized aerosol size distribution, and the aerosol number concentration.

The elements of the forward model vector $\mathbf{F}$ are in the most general case modeled values of light intensity and polarization at the top of the model atmosphere (at different wavelengths and in different viewing directions of the satellite). We will use the symbol E_i to refer to the modeled value of the ith Stokes parameter at the top of the atmosphere, at a given wavelength and viewing direction of the satellite measurement. E_i is called a radiative effect. The radiative effect E_i

can be extracted from the intensity vector field $\mathbf{I}$ (i.e. the solution of the radiative transfer equation (5.6)) with a suitable response vector function $\mathbf{R}_i$, via the inner product (see, for example, Marchuk (1964)):

$$E_i = \langle \mathbf{R}_i, \mathbf{I} \rangle. \tag{5.52}$$

Here, the inner product of two arbitrary vector functions $\mathbf{a}$ and $\mathbf{b}$ is defined by

$$\langle \mathbf{a}, \mathbf{b} \rangle = \int \mathrm{d}z \int \mathrm{d}\Omega \; \mathbf{a}^T(z, \mathbf{\Omega}) \mathbf{b}(z, \mathbf{\Omega}), \tag{5.53}$$

with integration over full solid angle and altitude range of the model atmosphere. The response functions $\mathbf{R}_i$ are given by

$$\mathbf{R}_i(z, \mathbf{\Omega}) = \delta(z - z_{\mathrm{top}}) \delta(\mathbf{\Omega} - \mathbf{\Omega}_v) \mathbf{e}_i, \tag{5.54}$$

where $\mathbf{e}_i$ is the unity vector in the direction of the ith component of the intensity vector, and $\mathbf{\Omega}_v = (\mu_v, \varphi_v)$ denotes the viewing direction of the instrument. In this context the response function formalism may seem somewhat awkward, but it is essential for a proper presentation of the adjoint formulation of radiative transfer, which will be described in section 5.4.1.

The requested derivatives of the elements of the forward model vector $\mathbf{F}$ in Eq. (5.2) can be expressed by the corresponding derivatives of the radiative effects E_i. Thus, the derivatives that we need to calculate are the derivatives $\partial E_i / \partial x_k$. These derivatives can be written as

$$\frac{\partial E_i}{\partial x_k} = \sum_{j=1}^{6} \sum_{l=0}^{L} \frac{\partial E_i}{\partial \alpha_j^l} \frac{\partial \alpha_j^l}{\partial x_k} + \frac{\partial E_i}{\partial K_{\mathrm{ext}}} \frac{\partial K_{\mathrm{ext}}}{\partial x_k} + \frac{\partial E_i}{\partial K_{\mathrm{sca}}} \frac{\partial K_{\mathrm{sca}}}{\partial x_k}. \tag{5.55}$$

Thus, the linearization corresponds to the calculation of two types of derivatives: (i) the derivatives $\partial E_i / \partial \alpha_i^j$, $\partial E_i / \partial K_{\mathrm{ext}}$, $\partial E_i / \partial K_{\mathrm{sca}}$ and (ii) the derivatives $\partial \alpha_j^l / \partial x_k$ $\partial K_{\mathrm{ext}} / \partial x_k$, and $\partial K_{\mathrm{sca}} / \partial x_k$. In the following we present an analytical approach to calculate these derivatives.

5.4.1 Linearization of radiative transfer

5.4.1.1 Forward-adjoint perturbation theory

For the linearization of radiative transfer with respect to the optical input parameters of the radiative transfer equation we will employ the forward-adjoint perturbation theory. Here, the adjoint formulation of radiative transfer is of essential importance. The transport operator adjoint to $\hat{\mathbf{L}}$, which is called $\hat{\mathbf{L}}^\dagger$, is defined by requiring that (see, for example, Marchuk (1964); Box et al. (1988))

$$\langle \mathbf{I}_2, \hat{\mathbf{L}} \mathbf{I}_1 \rangle = \langle \hat{\mathbf{L}}^\dagger \mathbf{I}_2, \mathbf{I}_1 \rangle \tag{5.56}$$

for arbitrary vector functions $\mathbf{I}_1$ and $\mathbf{I}_2$. The adjoint vector field $\mathbf{I}^\dagger$ is the solution of the adjoint transport equation

$$\hat{\mathbf{L}}^{\dagger}\mathbf{I}^{\dagger} = \mathbf{S}^{\dagger} \tag{5.57}$$

with any suitable adjoint source $\mathbf{S}^{\dagger}$. The adjoint operator $\hat{\mathbf{L}}^{\dagger}$ is given by (Marchuk, 1964; Carter et al., 1978)

$$\begin{aligned}\hat{\mathbf{L}}^{\dagger} \;=\; & \int_{4\pi} \mathrm{d}\tilde{\Omega} \Bigg\{ \Big[-\mu\frac{\partial}{\partial z} + K_{\mathrm{ext}}(z) \Big] \delta(\mathbf{\Omega} - \tilde{\mathbf{\Omega}})\, \mathbf{E} \\ & - \frac{K_{\mathrm{sca}}(z)}{4\pi} \mathbf{Z}^{T}(z, \mathbf{\Omega}, \tilde{\mathbf{\Omega}}) - \delta(z)\Theta(-\mu)|\mu| \mathbf{R}_s^{T}(\mathbf{\Omega}, \tilde{\mathbf{\Omega}})\, \Theta(\tilde{\mu})|\tilde{\mu}| \Bigg\}. \end{aligned} \tag{5.58}$$

The inclusion of the surface reflection term (last term on the right-hand side) is discussed by Ustinov (2001) and Landgraf et al. (2002). We see that compared to the forward operator $\hat{\mathbf{L}}$ the adjoint operator $\hat{\mathbf{L}}^{\dagger}$ has a different sign in the first term, and the phase matrix $\mathbf{Z}(z, \tilde{\mathbf{\Omega}}, \mathbf{\Omega})$ and the surface reflection matrix $\mathbf{R}_s(\tilde{\mathbf{\Omega}}, \mathbf{\Omega})$ are replaced by $\mathbf{Z}^{T}(z, \mathbf{\Omega}, \tilde{\mathbf{\Omega}})$, and $\mathbf{R}_s^{T}(\mathbf{\Omega}, \tilde{\mathbf{\Omega}})$, respectively. The adjoint vector field $\mathbf{I}^{\dagger}$ has to fulfill the boundary conditions (Box et al., 1988)

$$\begin{aligned} \mathbf{I}^{\dagger}(z_{\mathrm{top}}, \mathbf{\Omega}) &= [0,0,0,0]^{T} \qquad \text{for } \mu > 0, \\ \mathbf{I}^{\dagger}(0, \mathbf{\Omega}) &= [0,0,0,0]^{T} \qquad \text{for } \mu < 0. \end{aligned} \tag{5.59}$$

The forward radiative transfer equation (5.6) and the adjoint transport equation (5.57) do not describe two independent problems. The solutions $\mathbf{I}$ and $\mathbf{I}^{\dagger}$ are linked by the relation

$$\langle \mathbf{S}^{\dagger}, \mathbf{I} \rangle = \langle \mathbf{I}^{\dagger}, \mathbf{S} \rangle, \tag{5.60}$$

which can be derived in a straightforward fashion using Eqs (5.6), (5.56), and (5.57). We now take the response vector function $\mathbf{R}_i$ of Eq. (5.54) as the adjoint source $\mathbf{S}^{\dagger}$. In this particular case, the left-hand side of Eq. (5.60) represents the definition of the radiative effect E_i (see Eq. (5.52)). Thus we see from Eq. (5.60) that there are two ways of computing the radiative effect E_i. The first is the forward approach: solve the radiative transfer equation (5.6) and take the inner product of the response function $\mathbf{R}_i$ with the solution $\mathbf{I}$. The second is the adjoint approach: solve the adjoint transport equation (5.57) for the adjoint source $\mathbf{S}^{\dagger} = \mathbf{R}_i$ and take the inner product of its solution $\mathbf{I}^{\dagger}$ with the radiation source $\mathbf{S}$. Now also the physical meaning of the adjoint field becomes clear. Namely, the value of the adjoint field at a given altitude z_s and in a certain direction $\mathbf{\Omega}_s$ gives the effect of a point source $\delta(z - z_s, \mathbf{\Omega} - \mathbf{\Omega}_s)$ on the radiative effect E_i. In other words, the adjoint field gives us the importance of a radiation source anywhere in the atmosphere for the radiative effect E_i (Lewins, 1965). Thus, if the adjoint vector field is known the radiative effect E_i can be calculated for any radiation source via Eq. (5.60).

Let us consider an atmosphere with a set of optical parameters (α_j^l, K_{ext} and K_{sca}) contained in the vector $\mathbf{g}_o$. We call this atmosphere the unperturbed atmosphere. We denote the corresponding vector intensity field by $\mathbf{I}_o$, and the adjoint field corresponding to the adjoint source $\mathbf{R}_i$ by $\mathbf{I}_o^{\dagger}(\mathbf{R}_i)$. We also consider a perturbed atmosphere with a vector of optical parameters $\mathbf{g} = \mathbf{g}_o + \Delta\mathbf{g}$, where the optical parameters are perturbed in one layer of the model atmosphere. The radiative effect E_i for the perturbed atmosphere is given by (Marchuk, 1964)

$$E_i(\mathbf{g}) = E_i(\mathbf{g}_o) - \langle \mathbf{I}_o^\dagger(\mathbf{R}_i), \Delta\hat{\mathbf{L}}\,\mathbf{I}_o\rangle + \mathcal{O}(\Delta\mathbf{g}^2), \tag{5.61}$$

where $\mathcal{O}(\Delta\mathbf{g}^2)$ denotes second and higher order terms. The change $\Delta\hat{\mathbf{L}}$ in the radiative transfer operator $\hat{\mathbf{L}}$ caused by the perturbation $\Delta\mathbf{g}$ can be written as:

$$\Delta\hat{\mathbf{L}} = \sum_{k=1}^{K} \Delta g_k\, \Delta\hat{\mathbf{L}}_k, \tag{5.62}$$

where $\Delta\hat{\mathbf{L}}_k$ is the the change in $\Delta\hat{\mathbf{L}}$ per unit in parameter g_k, and K is the total number of optical parameters. The explicit form of $\Delta\hat{\mathbf{L}}_k$ follows from the definition of the transport operator $\hat{\mathbf{L}}$ (5.7). Substitution of Eq. (5.62) in Eq. (5.61) and comparison with a Taylor expansion yields the requested derivatives of the radiative effect E_i with respect to the optical parameters g_k:

$$\frac{\partial E_i}{\partial g_k} = -\langle \mathbf{I}_o^\dagger(\mathbf{R}_i), \Delta\hat{\mathbf{L}}_k\,\mathbf{I}_o\rangle. \tag{5.63}$$

So, in order to calculate the requested derivative the intensity vector field $\mathbf{I}_o$ is required as well as the adjoint fields $\mathbf{I}_o^\dagger$ for the adjoint sources $\mathbf{R}_i$ with $i = 1, \ldots, 4$.

5.4.1.2 Transformation to pseudo-forward problem

The adjoint field can be calculated with the same radiative transfer model as the forward intensity field, because the adjoint transport equation (5.57) may be transformed to a pseudo-forward problem. For this purpose we consider the vector function

$$\boldsymbol{\Psi}(z, \boldsymbol{\Omega}) = \mathbf{I}^\dagger(z, -\boldsymbol{\Omega}). \tag{5.64}$$

With substitution of Eq. (5.64) in Eq. (5.57), and with the symmetry relation of the scattering phase matrix (Hovenier, 1969)

$$\mathbf{Z}^T(z, -\boldsymbol{\Omega}, -\tilde{\boldsymbol{\Omega}}) = \mathbf{Q}\mathbf{Z}(z, \tilde{\boldsymbol{\Omega}}, \boldsymbol{\Omega})\mathbf{Q}, \tag{5.65}$$

with

$$\mathbf{Q} = \mathrm{diag}[1, 1, 1, -1], \tag{5.66}$$

and a similar relation for the surface reflection matrix $\mathbf{R}_s$, the adjoint transport equation transforms to a pseudo-forward equation

$$\hat{\mathbf{L}}_{\boldsymbol{\Psi}}\boldsymbol{\Psi} = \mathbf{S}_\Psi, \tag{5.67}$$

where

$$\mathbf{S}_\Psi(z, \boldsymbol{\Omega}) = \mathbf{R}_i(z, -\boldsymbol{\Omega}). \tag{5.68}$$

Here, the transport operator $\hat{\mathbf{L}}_{\boldsymbol{\Psi}}$ is the same as $\hat{\mathbf{L}}$ defined in Eq. (5.7), except that $\mathbf{Z}(z, \tilde{\boldsymbol{\Omega}}, \boldsymbol{\Omega})$ is replaced by $\mathbf{Q}\mathbf{Z}(z, \tilde{\boldsymbol{\Omega}}, \boldsymbol{\Omega})\mathbf{Q}$ and $\mathbf{R}_s(z, \tilde{\boldsymbol{\Omega}}, \boldsymbol{\Omega})$ is replaced by

$\mathbf{QR}_s(z, \tilde{\mathbf{\Omega}}, \mathbf{\Omega})\mathbf{Q}$. According to Eqs (5.64) and (5.59), $\mathbf{\Psi}$ has to fulfill the same boundary conditions as $\mathbf{I}$ in Eq. (5.10).

For the pseudo-forward problem a Fourier expansion can be performed as described in section 5.2. However, here the Fourier coefficients $\mathbf{S}_{\mathbf{\Psi}}^{\pm m}$ of the pseudo-forward source $\mathbf{S}_{\Psi}(z, \mathbf{\Omega}) = \mathbf{R}_i(z, -\mathbf{\Omega})$ depend on the index i, indicating the radiative effect E_i under consideration. For $i = 1, 2$, i.e. for the calculation of the derivatives of I and Q, we obtain

$$\begin{aligned} \mathbf{S}_{\mathbf{\Psi}}^{+m}(z, \mu) &= \frac{1}{2\pi}\, \delta(z - z_{top})\, \delta(\mu + \mu_v)\, \mathbf{e}_i, \\ \mathbf{S}_{\mathbf{\Psi}}^{-m}(z, \mu) &= [0, 0, 0, 0]^T. \end{aligned} \tag{5.69}$$

Hence, for the corresponding pseudo-forward problems we obtain a Fourier expansion of $\mathbf{\Psi}$ containing terms of $\mathbf{\Psi}^{+m}$ only. For $i = 3, 4$, i.e. for the calculation of the derivatives of U and V, we obtain

$$\begin{aligned} \mathbf{S}_{\mathbf{\Psi}}^{+m}(z, \mu) &= [0, 0, 0, 0]^T, \\ \mathbf{S}_{\mathbf{\Psi}}^{-m}(z, \mu) &= \frac{1}{2\pi}\, \delta(z - z_{top})\, \delta(\mu + \mu_v)\, \mathbf{e}_i. \end{aligned} \tag{5.70}$$

Hence, for the corresponding pseudo-forward problems we obtain a Fourier expansion of $\mathbf{\Psi}$ containing terms of $\mathbf{\Psi}^{-m}$ only.

5.4.1.3 Calculation of the derivatives

In the following we will work out the expressions for the derivatives with respect to the expansion coefficients α_j^l, the scattering coefficient K_{sca}, and the extinction coefficient K_{ext}. Hereto, we write instead of Eq. (5.62)

$$\Delta\hat{\mathbf{L}} = \sum_{j=1}^{6} \sum_{l=0}^{L} \Delta\alpha_j^l\, \Delta\hat{\mathbf{L}}_j^l \;+\; \Delta K_{\mathrm{sca}} \Delta\hat{\mathbf{L}}_{\mathrm{sca}} \;+\; \Delta K_{\mathrm{ext}} \Delta\hat{\mathbf{L}}_{\mathrm{ext}}, \tag{5.71}$$

where

$$\begin{aligned} \Delta\hat{\mathbf{L}}_j^l &= \frac{\beta^s}{4\pi} \int_{4\pi} \mathrm{d}\tilde{\Omega} \frac{\partial \mathbf{Z}(z, \tilde{\mathbf{\Omega}}, \mathbf{\Omega})}{\partial \alpha_j^l}, && (5.72) \\ \Delta\hat{\mathbf{L}}_{\mathrm{sca}} &= \frac{1}{4\pi} \int_{4\pi} \mathrm{d}\tilde{\Omega} \mathbf{Z}(z, \tilde{\mathbf{\Omega}}, \mathbf{\Omega}), && (5.73) \\ \Delta\hat{\mathbf{L}}_{\mathrm{ext}} &= \int_{4\pi} \mathrm{d}\tilde{\Omega}\; \delta(\mathbf{\Omega} - \tilde{\mathbf{\Omega}})\, \mathbf{E}. && (5.74) \end{aligned}$$

In order to obtain expressions for the derivatives with respect to α_j^l, K_{sca}, and K_{ext}, we substitute $\Delta\hat{\mathbf{L}}_j^l$, $\Delta\hat{\mathbf{L}}_{\mathrm{sca}}$, and $\Delta\hat{\mathbf{L}}_{\mathrm{ext}}$ in Eq. (5.63), respectively. Additionally, we use the Fourier expansion of $\mathbf{I}$, $\mathbf{\Psi}$, and $\mathbf{Z}$, and evaluate the integrals over azimuth angle. We then obtain expressions in the form of cosine- and sine expansions which have a similar form for the three types of derivatives. For the radiative effects E_i with $i = 1, 2$ (i.e. corresponding to Stokes parameters I and Q, respectively), the derivatives are given by a cosine expansion:

$$\frac{\partial E_i}{\partial g_k} = -\sum_{m=0}^{\infty} (2 - \delta_{m0}) \, \cos m(\phi_v - \phi_0) \, K_i^{+m}(g_k). \tag{5.75}$$

For the radiative effects E_i with $i = 3, 4$ (i.e. corresponding to Stokes parameters U and V, respectively) the derivatives are given by a sines expansion:

$$\frac{\partial E_i}{\partial g_k} = -\sum_{m=0}^{\infty} (2 - \delta_{m0}) \, \sin m(\phi_v - \phi_0) \, K_i^{-m}(g_k). \tag{5.76}$$

The specific integral kernels for α_j^l, K_{sca}, and K_{ext} are determined by $\Delta\hat{\mathbf{L}}_j^l$, $\Delta\hat{\mathbf{L}}_{\mathrm{sca}}$, and $\Delta\hat{\mathbf{L}}_{\mathrm{ext}}$, respectively. For the derivative of the radiative effect E_i with respect to the expansion coefficient α_j^l, we obtain for the integral kernel

$$\begin{aligned} K_i^{\pm m}(\alpha_j^l) &= \frac{\pi}{4} \int_{z_{\mathrm{bot}}}^{z_{\mathrm{top}}} \mathrm{d}z \; K_{\mathrm{sca}}(z) \\ & \int_{-1}^{1} \int_{-1}^{1} \mathrm{d}\mu \, \mathrm{d}\tilde{\mu} \; \mathbf{\Psi}_o^{\pm mT}(\mathbf{R}_i, z, -\mu) \, \boldsymbol{\Lambda} \, \dot{\mathbf{Z}}^m(z, \tilde{\mu}, \mu) \, \mathbf{I}_o^{+m}(z, \tilde{\mu}), \end{aligned} \tag{5.77}$$

where $\boldsymbol{\Lambda} = \mathrm{diag}\,[1, 1, -1, -1]$, and the derivative $\dot{\mathbf{Z}}^m = \partial \mathbf{Z}^{\mathbf{m}} / \partial \alpha_j^l$ is given by

$$\dot{\mathbf{Z}}^m(z, \tilde{\mu}, \mu) = (-1)^m \; \mathbf{P}_m^l(-\mu) \, \mathbf{H}_j \; \mathbf{P}_m^l(-\tilde{\mu}). \tag{5.78}$$

Here, the matrix $\mathbf{H}_j$ has the same structure as the expansion coefficient matrix (5.23) and is given by

$$\mathbf{H}_j = \begin{pmatrix} \delta_{j1} & \delta_{j5} & 0 & 0 \\ \delta_{j5} & \delta_{j2} & 0 & 0 \\ 0 & 0 & \delta_{j3} & \delta_{j6} \\ 0 & 0 & -\delta_{j6} & \delta_{j4} \end{pmatrix}, \tag{5.79}$$

where δ is the Kronecker delta.

The integral kernel corresponding to the derivative of the radiative effect E_i with respect to K_{sca} is given by:

$$K_i^{\pm m}(K_{\mathrm{sca}}) = \frac{\pi}{4} \int_{z_{\mathrm{bot}}}^{z_{\mathrm{top}}} \mathrm{d}z \int_{-1}^{1} \int_{-1}^{1} \mathrm{d}\mu \, \mathrm{d}\tilde{\mu} \; \mathbf{\Psi}_o^{\pm mT}(\mathbf{R}_i, z, -\mu) \, \boldsymbol{\Lambda} \, \mathbf{Z}^m(z, \tilde{\mu}, \mu) \, \mathbf{I}_o^{+m}(z, \tilde{\mu}), \tag{5.80}$$

and the integral kernel corresponding to the derivative of the radiative effect E_i with respect to K_{ext} has the following form:

$$K_i^{\pm m}(K_{\mathrm{ext}}) = 2\pi \int_{z_{\mathrm{bot}}}^{z_{\mathrm{top}}} \mathrm{d}z \int_{-1}^{1} \mathrm{d}\mu \; \mathbf{\Psi}_o^{\pm mT}(\mathbf{R}_i, z, -\mu) \, \boldsymbol{\Lambda} \, \mathbf{I}_o^{+m}(z, \mu). \tag{5.81}$$

Equations (5.75), (5.76), together with (5.77)–(5.81) provide analytical expressions for the derivatives of the radiative effect E_i with respect to the expansion coefficients α_j^l, the scattering coefficient K_{sca}, and the extinction coefficient K_{ext}, respectively. Thus, to calculate these derivatives one needs to solve the forward radiative transfer equation (5.6) and the adjoint transport equation (5.57) for the sources $\mathbf{S}_{\boldsymbol{\Psi}}(z,\boldsymbol{\Omega}) = \mathbf{R}_i(z,-\boldsymbol{\Omega})$, with $i = 1, 4$. These fields can be determined by any vector radiative transfer model that calculates the internal radiation in the atmosphere, such as the doubling and adding model of Stammes et al. (1989), the discrete ordinate model VDISORT of Schulz et al. (1999) and the Gauss–Seidel model of Hasekamp and Landgraf (2002). The latter model is used for all numerical simulations in this chapter. The corresponding integral kernels of equations (5.77)–(5.81) are worked out by Landgraf et al. (2004) for this model.

5.4.2 Linearization of Mie theory

The derivatives of the optical input parameters of the radiative transfer equation with respect to the different elements x_k of the state vector can be found from the corresponding derivatives of the optical aerosol parameters (see Eqs (5.31)–(5.33)):

$$
\begin{aligned}
\frac{\partial K_{\text{ext}}}{\partial x_k} &= \frac{\partial K_{\text{ext}}^a}{\partial x_k}, && (5.82)\\
\frac{\partial K_{\text{sca}}}{\partial x_k} &= \frac{\partial K_{\text{sca}}^a}{\partial x_k}, && (5.83)\\
\frac{\partial \alpha}{\partial x_k} &= \frac{K_{\text{sca}}^a}{K_{\text{sca}}}\frac{\partial \alpha^a}{\partial x_k} + \frac{\alpha^a}{K_{\text{sca}}}\frac{\partial K_{\text{sca}}^a}{\partial x_k} - \frac{K_{\text{sca}}^a\,\alpha^a}{(K_{\text{sca}})^2}\frac{\partial K_{\text{sca}}^a}{\partial x_k} - \frac{K_{\text{sca}}^r\,\alpha^r}{(K_{\text{sca}})^2}\frac{\partial K_{\text{sca}}^a}{\partial x_k}, && (5.84)
\end{aligned}
$$

where we omitted the indices for the expansion coefficients α_j^l. In this subsection we will derive expressions for the requested derivatives $\partial\alpha^a/\partial x_k$, $\partial K_{\text{ext}}^a/\partial x_k$, and $\partial K_{\text{sca}}^a/\partial x_k$. For notational convenience, we will omit the superscript a for the expansion coefficients in the remainder of this subsection.

In order to calculate the derivatives of the expansion coefficients α_j^l with respect to the real and imaginary part of the refractive index, m_r and m_i, respectively, we first need to calculate the corresponding derivatives of the elements f_j of the transformation matrix $\boldsymbol{\mathcal{F}}$ in Eq. (5.34). These derivatives are expressed via the derivatives of S_1 and S_2 (see Eqs (5.35)–(5.40)):

$$[f_1]' = \frac{1}{2}\left(S_1[S_1^*]' + [S_1]' S_1^* + S_2[S_2^*]' + [S_2]' S_2^*\right), \tag{5.85}$$

$$[f_2]' = [f_1]', \tag{5.86}$$

$$[f_3]' = \frac{1}{2}\left(S_1[S_2^*]' + [S_1]' S_2^* + S_2[S_1^*]' + [S_2]' S_1^*\right), \tag{5.87}$$

$$[f_4]' = [f_3]', \tag{5.88}$$

$$[f_5]' = \frac{1}{2}\big(S_1[S_1^*]' + [S_1]' S_1^* - S_2[S_2^*]' + [S_2]' S_2^*\big), \tag{5.89}$$

$$[f_6]' = \frac{i}{2}\big(S_1[S_2^*]' + [S_1]' S_2^* - S_2[S_1^*]' + [S_2]' S_1^*\big). \tag{5.90}$$

Here, the prime denotes the derivative with respect to either m_r or m_i. The derivatives of S_1 and S_2 are given by

$$[S_1]' = \sum_{n=1}^{\infty} \frac{2n+1}{n(n+1)} \left([a_n]' \pi_n + [b_n]' \tau_n\right), \tag{5.91}$$

$$[S_2]' = \sum_{n=1}^{\infty} \frac{2n+1}{n(n+1)} \left([b_n]' \pi_n + [a_n]' \tau_n\right). \tag{5.92}$$

The derivatives of the scattering and extinction coefficients with respect to m_r and m_i follow from the corresponding derivatives of the scattering and extinction cross-sections. These derivatives also depend on the derivatives of a_n and b_n:

$$[C_{\text{sca}}]' = \frac{4\pi}{k^2} \sum_{n=1}^{\infty} (2n+1)\big(a_n[a_n^*]' + [a_n]' a_n^* + b_n[b_n^*]' + [b_n]' b_n^*\big), \tag{5.93}$$

$$[C_{\text{ext}}]' = \frac{2\pi}{k^2} \sum_{n=1}^{\infty} (2n+1)\text{Re}([a_n]' + [b_n]'). \tag{5.94}$$

Thus, all derivatives with respect to m_r and m_i depend on the corresponding derivatives of the Mie coefficients a_n and b_n, for which analytical expressions are given in Appendix A.

The derivatives of f_j, C_{sca}, and C_{ext} with respect to m_r and m_i given above all correspond to a single sphere with a given radius. The derivatives for an ensemble of particles with a given size distribution can be easily obtained via integration over the size distribution as in Eq. (5.50). The derivatives of elements p_j of the scattering phase matrix (5.5) for the size distribution can then be calculated using the derivatives of $\bar{f}_j$ and $\bar{C}_{\text{sca}}$, viz.

$$[p_j]' = \frac{4\pi}{k^2} \left(\frac{[\bar{f}_j]'}{\bar{C}_{\text{sca}}} - \frac{\bar{f}_j [C_{\text{sca}}]'}{\bar{C}_{\text{sca}}^2} \right). \tag{5.95}$$

The derivatives of the expansion coefficients α_j^l can be calculated from Eqs (5.24)–(5.29), replacing α_i^j and p_j by their corresponding derivatives.

The other type of derivatives that are needed are the derivatives with respect to the elements n_i of the discretized size distribution (see Eq. (5.50)). These derivatives can be calculated in a straightforward manner. The derivative of an averaged element $\bar{f}_j$ of the transformation matrix $\boldsymbol{\mathcal{F}}$ with respect to n_i is given by:

$$\frac{\partial \bar{f}_j}{\partial n_i} = f_j(r_i)\, \Delta r_i, \tag{5.96}$$

and similar expressions hold for the derivatives of $\bar{C}_{\text{sca}}$ and $\bar{C}_{\text{ext}}$ with respect to n_i. Using the derivatives of $\bar{f}_j$ and $\bar{C}_{\text{sca}}$ with respect to n_i, the derivatives of the elements p_j of the scattering phase matrix (5.5) with respect to n_i be calculated using Eq. (5.95). The corresponding derivatives of the aerosol expansion coefficients α_i^j can be calculated from Eqs (5.24)–(5.29), replacing α_i^j and p_j by their corresponding derivatives.

For aerosol retrieval it is often useful to describe the size distribution by a limited number of parameters, for example the effective radius r_{eff} and the effective variance v_{eff} (see Appendix B) of a prescribed size distribution (e.g. a lognormal distribution). For such retrievals one needs to calculate the derivatives with respect to these parameters. For an averaged parameter $\bar{g}$ this derivative is given by

$$\frac{\partial \bar{g}}{\partial r_{\text{eff}}} = \sum_{i=1}^{N} \frac{\partial \bar{n}_i}{\partial r_{\text{eff}}} \frac{\partial \bar{g}}{\partial \bar{n}_i}, \tag{5.97}$$

and a similar expression holds for the derivative with respect to v_{eff}.

5.5 Numerical implementation and results

The linearization approach described in section 5.4 has been implemented in the Gauss-Seidel vector radiative transfer model described by Landgraf et al. (2002) and Hasekamp and Landgraf (2002), combined with the Mie scattering algorithm of de Rooij and van der Stap (1984). Expressions for the integral kernels of section 5.4.1 can be found in the paper of Landgraf et al. (2004) for our Gauss–Seidel radiative transfer model.

All radiative transfer calculations in this section were performed for a model atmosphere that includes Rayleigh scattering and scattering and absorption by homogeneous spherical aerosol particles. All aerosols were homogeneously distributed over the lowest 2 km of the atmosphere. We used a bi-modal lognormal aerosol size distribution, with a mode containing small particles referred to as the small mode and a mode containing large particles referred to as the large mode. For this model atmosphere the aerosols are characterized by ten parameters, i.e. five per mode of the size distribution. These parameters are: per mode the effective radius r_{eff}, the effective variance v_{eff}, the column integrated aerosol number concentration N, and the real and imaginary part of the refractive index m.

Figures 5.1 and 5.2 show the derivatives of Stokes parameters I and Q at the top of the atmosphere with respect to the logarithm of the effective radius, real refractive index, and aerosol loading of the two size modes, as a function of viewing zenith angle for a solar zenith angle of 40^o, for a wavelength of 350 nm and 800 nm, respectively. The relative azimuth angle $\phi_o - \phi_v = 180^o$ for negative viewing zenith angles and $\phi_o - \phi_v = 0^o$ for positive viewing zenith angles. For these geometries the Stokes parameters U and V are equal to zero, so Stokes parameter Q fully describes the polarization of the backscattered light at the top of the atmosphere. The derivative with respect to the logarithm of a given

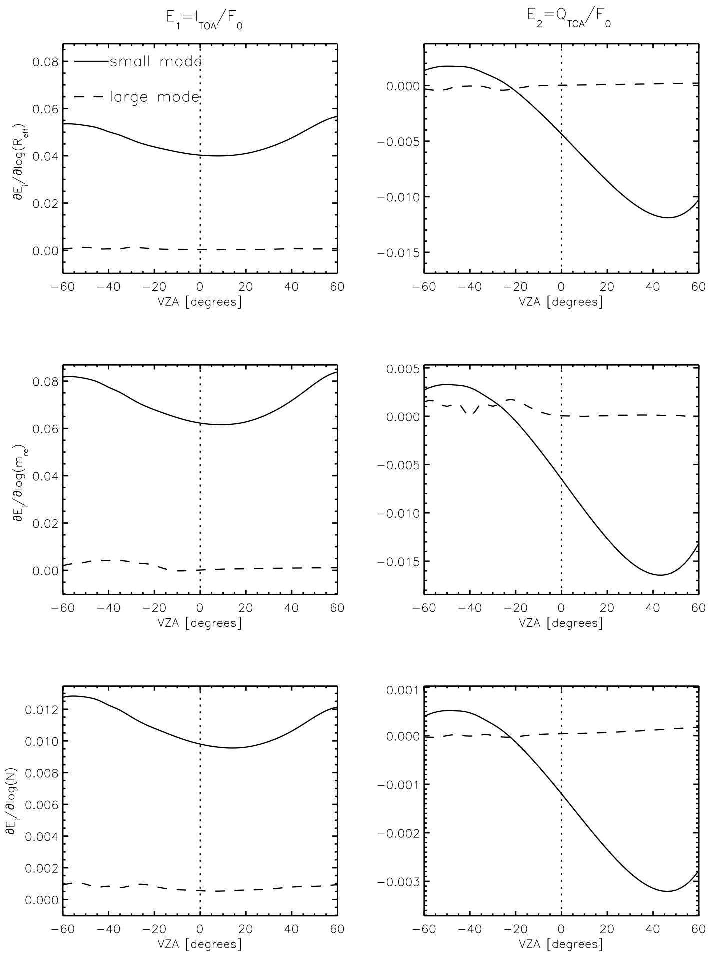

Fig. 5.1. Derivatives of the Stokes parameters I (left panels) and Q (right panels) at the top of the atmosphere (TOA), with respect to the logarithm of: (upper panels) r_{eff}, (middle panels) m_r, and (lower panels) the column integrated aerosol number concentration N. The solid lines correspond to parameters of the small mode and the dashed lines correspond to parameters of the large mode. The derivatives are shown as a function of viewing zenith angle (VZA), where $VZA < 0$ refer to the relative azimuth angle $\phi_o - \phi_v = 180^o$, and $VZA > 0$ refer to $\phi_o - \phi_v = 0^o$. The solar zenith angle is 40^o, and the calculation is performed for a wavelength of 350 nm. The solar and viewing zenith angles are defined as the smallest angle between the zenith direction and the solar and viewing direction, respectively. The range -60^o to 60^o of viewing zenith angles used in this figure corresponds to a horizon-to-horizon scan from a satellite at approximately 800 km. The internal radiation field was discretized in 16 Gaussian streams. A bimodal aerosol size distribution was used with $r_{\text{eff}} = 0.05$ for the small mode, $r_{\text{eff}} = 0.75$ for the large mode, $v_{\text{eff}} = 0.2$ for both modes, $m_r = 1.45$, $m_i = -0.0045$. The optical thickness at 550 nm is 0.15 with equal contribution from the small and the large mode. The model atmosphere is bounded below by a black surface.

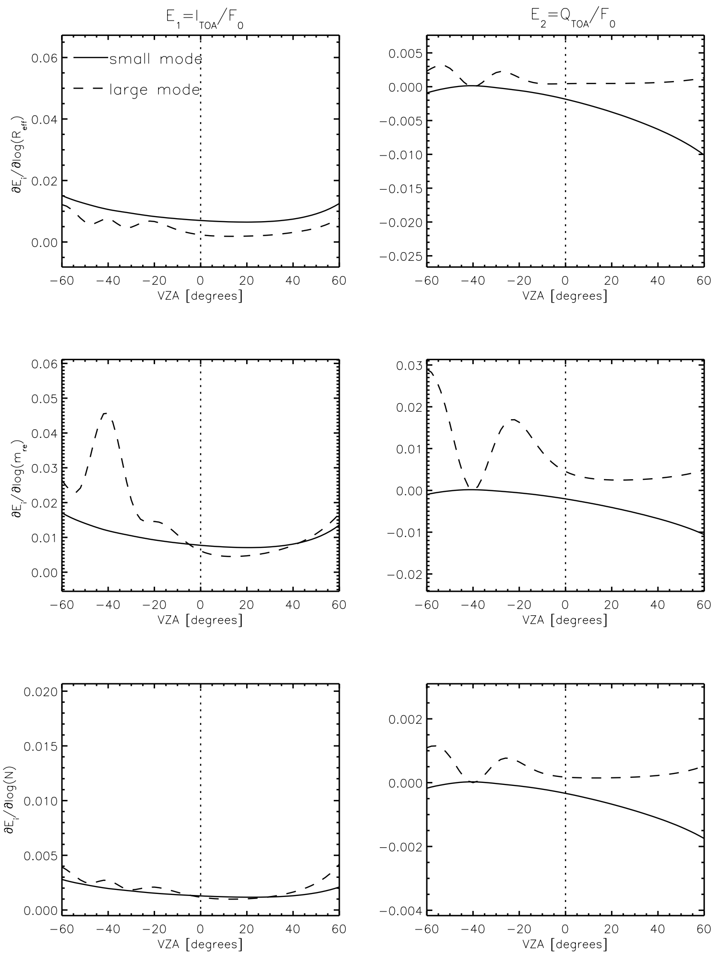

Fig. 5.2. Same as Fig. 5.1 but for a wavelength of 800 nm.

aerosol parameter is a measure for the sensitivity of I and Q to a relative change in this aerosol parameter, which is a convenient quantity in order to compare the sensitivities to the different parameters.

The angular dependence of the derivatives is caused by the angular dependence of the following effects: (i) The derivatives of the relevant elements of the aerosol scattering phase matrix with respect to the different aerosol parameters. (ii) The light path inside the aerosol layer, which increases with viewing zenith angle. This causes an increase in sensitivity up to a certain viewing angle because an increasing fraction of the light is scattered by aerosol particles. However, if the viewing angle becomes too large this effect causes a decrease in sensitivity because of increasing extinction within the aerosol layer along the line of sight. (iii) Multiple scattering effects, which in general smear out the angular effects of the aerosol scattering phase matrix.

From Fig. 5.1 it follows that at 350 nm the sensitivities of both I and Q with respect to the parameters of the small mode are much larger than the corresponding sensitivities with respect to the parameters of the large mode, which are in general negligibly small at 350 nm. The angular dependence of the derivatives of the intensity in Fig. 5.1 is relatively weak which is caused by the weak angular dependence of the corresponding derivatives of element (1,1) of the scattering phase matrix (except for the forward scattering direction, which is not shown in Fig. 5.1). The derivatives with respect to the different aerosol parameters of element (2,1) of the scattering phase matrix have a larger angular dependence, which is the most dominant effect in the right panels of Fig. 5.1. Also multiple scattering effects can be seen here, because both the Rayleigh and aerosol scattering optical thickness are relatively large at 350 nm. For example, in the backward scattering direction element (2,1) of the scattering phase matrix is zero independent of the aerosol properties, i.e. it is insensitive with respect to aerosol properties. However, the sensitivity of Stokes parameter Q is not zero in the backward single scattering direction (viewing angle = $-40°$), because the sensitivity is influenced by aerosol scattering in all directions via multiple scattering.

At 800 nm (Fig. 5.2) the derivatives of both I and Q with respect to the parameters of the large mode, are significantly larger than at 350 nm, while the derivatives with respect to the parameters if the small mode are much smaller. The angular dependence of the derivatives of the phase matrix plays the most important role at 800 nm, especially for the derivatives of Stokes parameter Q. Here, the strong angular variation in the derivatives of Q around the single scattering backward direction ($-40°$) is also present in the sensitivity of the (2,1)-element of the aerosol scattering phase matrix. A similar, but weaker effect can be seen in the derivatives of the intensity with respect to parameters of the large mode. Another effect that can be seen in Fig. 5.2 is the slight increase in sensitivity towards larger (absolute values of) viewing zenith angle. This increase in sensitivity is caused by an enhanced light path inside the aerosol layer.

5.6 Retrieval method

In this section we will discuss how the linearized vector radiative transfer model can be incorporated in a retrieval algorithm for aerosol properties over the ocean. Here, we assume that the aerosol size distribution can be described by a bi-modal lognormal function, where each mode is characterized by the effective radius r_{eff}, the effective variance v_{eff} (see Appendix B) and the column integrated aerosol number concentration N. In what follows we use the superscripts l and s to refer to the small and large mode of the size distribution, respectively. Additionally, the complex refractive index $m = m_r + im_i$ is needed to characterize aerosols. Furthermore, we assume an altitude distribution with a constant aerosol density ρ_o in the lowest layer with height z_b of the atmosphere. Above that layer the aerosol density decreases with the fourth power in pressure p till a certain height z_t above which we assume no aerosols are located, i.e.

$$\begin{aligned} \rho(z) &= \rho_o && \text{for } z < z_b, \\ \rho(z) &= \rho_o(p(z)/p(z_b))^4 && \text{for } z_b < z < z_t, \\ \rho(z) &= 0 && \text{for } z > z_t. \end{aligned} \tag{5.98}$$

In the retrievals described here nine unknown aerosol parameters are considered. These are the effective radius r_{eff} of the small and large mode, the effective variance v_{eff} of the small and large mode, the column integrated aerosol number concentration N of the small and large mode, the real and imaginary part of the refractive index, and the height z_b of the layer where the bulk of the aerosols is located. Here, we assume that the wavelength dependence of the refractive index is known. For all retrieval simulations in this paper the model atmosphere is bounded below by a rough ocean surface (see Appendix B). Here, the oceanic pigment concentration is included as an additional parameter to be retrieved in addition to the nine aerosol parameters.

5.6.1 Inversion of linearized forward model

In this subsection we consider the inversion of the linearized forward model (5.2), assuming that the state vector $\mathbf{x}_n$ of the iteration step under consideration is close enough to the true state vector so that the linear approximation is valid. In this case the inversion of Eq. (5.2) provides the solution of our retrieval problem. Rearranging terms in Eq. (5.2) we obtain

$$\tilde{\mathbf{y}} = \mathbf{K}\,\mathbf{x} + \mathbf{e}_y, \tag{5.99}$$

with $\tilde{\mathbf{y}} = \mathbf{y} - \mathbf{F}(\mathbf{x}_\text{n}) + \mathbf{K}\mathbf{x}_\text{n}$. Here, $\mathbf{y}$ is the measurement vector, $\mathbf{F}(\mathbf{x}_\text{n})$ is the forward model vector for state vector $\mathbf{x}_\text{n}$, $\mathbf{K}$ is the Jacobian matrix defined by Eq. (5.3), and $\mathbf{e}_y$ is the measurement error vector.

For most types of satellite instruments the inversion of Eq. (5.99) represents an ill-posed problem. This means that many combinations of the state vector parameters fit the measurement almost equally well. As a result, the least-squares solution $\hat{\mathbf{x}}_\text{lsq}$ to our retrieval problem, viz.

$$\hat{\mathbf{x}}_\text{lsq} = \min_{\mathbf{x}} ||\mathbf{S}_y^{-\frac{1}{2}}(\mathbf{K}\mathbf{x} - \tilde{\mathbf{y}})||^2, \tag{5.100}$$

is overwhelmed by noise. In order to reduce the effect of noise, we use the Phillips–Tikhonov regularization method (Phillips, 1962; Tikhonov, 1963), which introduces a side constraint in addition to the minimization of the least-squares norm. As a side constraint we choose for our application the minimization of a weighted norm of the state vector, viz.

$$\hat{\mathbf{x}}_\text{reg} = \min_{\mathbf{x}} \left(||\mathbf{S}_y^{-\frac{1}{2}}(\mathbf{K}\mathbf{x} - \tilde{\mathbf{y}})||^2 + \gamma\, ||\mathbf{\Gamma}\mathbf{x}||^2 \right), \tag{5.101}$$

where $\mathbf{\Gamma}$ is a diagonal matrix that contains weighting factors for the different state vector elements in the side constraint, and the regularization parameter γ balances the two minimizations in Eq. (5.101). For each iteration step the solution $\hat{\mathbf{x}}_{\mathrm{reg,n+1}}$ in Eq. (5.101) can be written as a matrix equation:

$$\hat{\mathbf{x}}_{\mathrm{reg}} = \mathbf{D}\,\tilde{\mathbf{y}}, \tag{5.102}$$

where $\mathbf{D}$ is the contribution matrix defined by

$$\mathbf{D} = \left(\mathbf{K}^T\,\mathbf{S}_y^{-1}\,\mathbf{K} + \gamma\mathbf{\Gamma}\right)^{-1}\;\mathbf{K}^T\,\mathbf{S}_y^{-1}, \tag{5.103}$$

where the superscript T denotes the transposed matrix.

The rationale of minimizing the norm of the state vector as a side constraint in Eq. (5.101) is to reduce the effect of measurement noise on the solution. Since the norm of the state vector is a quantity that is very sensitive to noise contributions, these contributions are reduced using Eq. (5.101) instead of Eq. (5.100). Clearly, a good choice of γ is of crucial importance for the Phillips–Tikhonov solution. If γ is chosen too large, the noise contribution will be low but the least squares norm deviates significantly from its minimum value, i.e. the fit between forward model and measurement is poor. On the other hand, if γ is chosen too small the measurement is fitted well but the solution norm is high, i.e. the solution is overwhelmed by noise. Thus, γ should be chosen such that the two minimizations are well balanced. Such a value for γ can be found from the L-curve (Hansen and O'Leary, 1993). The L-curve is a parametric plot of the weighted least-squares norm $||\mathbf{S}_y^{-\frac{1}{2}}(\mathbf{K}(\mathbf{x}) - \tilde{\mathbf{y}})||$ and the weighted solution norm $||\mathbf{\Gamma}\mathbf{x}||$, with a characteristic L-shaped corner. The corner of the L-curve corresponds to the optimum value of the regularization parameter. A numerical stable and efficient method for determining the corner of the L-curve is given by Hansen (1992), who defines the corner of the L-curve as the point with maximum curvature, where the curvature is calculated analytically. Visual inspection of the L-curves of our retrievals showed that in all cases the method of Hansen (1992) provided a value for the regularization parameter that corresponds to the 'true' corner of the L-curve. An example of an L-curve with the corresponding curvature is given in Fig. 5.3 for aerosol retrieval from synthetic GOME-2 measurements of intensity and polarization.

Due to the inclusion of the side constraint, the state vector $\hat{\mathbf{x}}_{\mathrm{reg}}$ retrieved using Eq. (5.101) does not represent an estimate of the true state vector $\mathbf{x}_{\mathrm{true}}$, but its elements represent weighted averages of the elements of $\mathbf{x}_{\mathrm{true}}$. The relation between $\hat{\mathbf{x}}_{\mathrm{reg}}$ and $\mathbf{x}_{\mathrm{true}}$ is expressed by the averaging kernel $\mathbf{A}$ (Rodgers, 2000), viz.

$$\hat{\mathbf{x}}_{\mathrm{reg}} = \mathbf{A}\mathbf{x}_{\mathrm{true}} + \mathbf{e}_x. \tag{5.104}$$

Here, $\mathbf{e}_x$ represents the error in the state vector caused by measurement errors, and the averaging kernel is given by

$$\mathbf{A} = \frac{\partial\hat{\mathbf{x}}}{\partial\mathbf{x}_{\mathrm{true}}} = \frac{\partial\hat{\mathbf{x}}}{\partial\mathbf{x}_{\mathrm{true}}} = \left(\mathbf{K}^T\,\mathbf{S}_y^{-1}\,\mathbf{K} + \gamma\mathbf{\Gamma}\right)^{-1}\;\mathbf{K}^T\,\mathbf{S}_y^{-1}\,\mathbf{K} = \mathbf{D}\,\mathbf{K}. \tag{5.105}$$

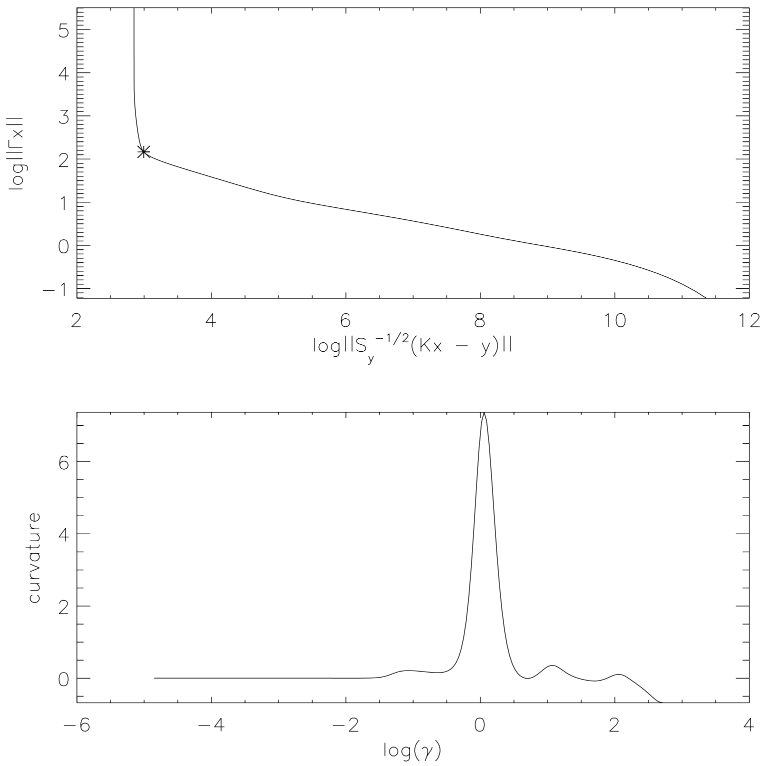

Fig. 5.3. L-curve (upper panel) and corresponding curvature (lower panel) for aerosol retrieval from synthetic GOME-2 measurements of intensity and polarization.

The matrix $\mathbf{A}$ is strongly related to the information content of the measurement $\mathbf{y}$, i.e. the closer $\mathbf{A}$ is to the unity matrix, the higher the information content. From the matrix $\mathbf{A}$ the Degrees of Freedom for Signal (DFS) can be derived (Rodgers, 2000), which indicates the number of independent pieces of information that is retrieved from the measurement:

$$\mathrm{DFS} = \mathrm{trace}\,(\mathbf{A}). \tag{5.106}$$

If $\mathbf{x}_{\mathrm{true}}$ would have represented a discretization of a continuous function, then the weighted averages contained in $\hat{\mathbf{x}}_{\mathrm{reg}}$ Eq. (5.104) would have a clear physical meaning, i.e. an estimate of $\mathbf{x}_{\mathrm{true}}$ at a reduced resolution. However, for our application the elements of $\hat{\mathbf{x}}_{\mathrm{reg}}$ represent weighted averages of different aerosol parameters, which have a limited physical meaning. Therefore, we include information from an *a priori* state vector $\mathbf{x}_a$ in the solution to make it a meaningful estimate of the $\mathbf{x}_{\mathrm{true}}$. Hereto, we add the term $(\mathbf{I} - \mathbf{A})\mathbf{x}_a$ to $\hat{\mathbf{x}}_{\mathrm{reg}}$ in order to obtain the final retrieval product $\hat{\mathbf{x}}$, viz.

$$\begin{aligned} \hat{\mathbf{x}} &= \hat{\mathbf{x}}_{\mathrm{reg}} + (\mathbf{I} - \mathbf{A})\mathbf{x}_a, \\ &= \mathbf{A}\mathbf{x}_{\mathrm{true}} + (\mathbf{I} - \mathbf{A})\mathbf{x}_a + \mathbf{e}_x. \end{aligned} \tag{5.107}$$

Thus, in Eq. (5.107) $\hat{\mathbf{x}}_{\text{reg}}$ represents an estimate of $\mathbf{A}\mathbf{x}_{\text{true}}$ and $(\mathbf{I}-\mathbf{A})\mathbf{x}_a$ represents an estimate of the part $(\mathbf{I}-\mathbf{A})\mathbf{x}_{\text{true}}$ of the true state vector that cannot be retrieved from the measurement. Here, the dependence of a retrieved element $\hat{x}_i$ of the state vector on its corresponding *a priori* value $x_{a,i}$ is given by

$$\frac{\partial \hat{x}_i}{\partial x_{a,i}} = 1 - a_{ii}, \tag{5.108}$$

where a_{ii} is element (i,i) of $\mathbf{A}$. An equation similar to Eq. (5.107) has been used by Rodgers and Connor (2003) to represent retrieval results with respect to a different *a priori* state vector than had been used in the retrieval. The reason that we first solve the minimization problem (5.101) and later add *a priori* information in Eq. (5.107), instead of directly including $\mathbf{x}_a$ in the side constraint of Eq. (5.101), is that in our approach the amount of information extracted from the measurement is independent of the *a priori* state vector $\mathbf{x}_a$. So, this approach is especially suited for characterizing the information content of satellite measurements.

The weighting factors in the matrix $\mathbf{\Gamma}$ are defined relative to the values of the corresponding state vector element for the iteration step under consideration. This makes the vector $\mathbf{\Gamma x}$ dimensionless. From Eq. (5.101) it can be seen that if the weighting factor for a certain parameter decreases while the other weighting factors are kept constant, more information about this parameter is obtained from the measurement. This means that the parameters with small relative weight are less dependent on the *a priori* information added in Eq. (5.107). So, if for certain state vector elements less reliable *a priori* information is available than for others, the relative weighting factors corresponding to these parameters should be set to small values. In this way the dependence on *a priori* information for the state vector elements with small relative weight becomes smaller while for the other parameters the dependence on *a priori* assumptions becomes larger, compared to the situation where all parameters have unity relative weight. For our application it may be expected that no reliable *a priori* information will be available for the aerosol columns of both modes, because these two parameters are highly variable. Therefore, the relative weighting factors corresponding to these two parameters are set to a very low value ϵ while the other factors get a unity relative weight. We found that for $\epsilon = 1 \times 10^{-8}$ the retrieved aerosol columns for both modes are virtually independent of their *a priori* values.

From Eq. (5.107) it is clear that the retrieved state vector $\hat{\mathbf{x}}$ is affected by errors in the *a priori* state vector $\mathbf{x}_a$. The error on $\hat{\mathbf{x}}$ caused by an error on $\mathbf{x}_a$ is called the regularization error (called smoothing error by Rodgers (2000)). The regularization error covariance matrix $\mathbf{S}_r$ is given by

$$\mathbf{S}_r = (\mathbf{I}-\mathbf{A})\,\mathbf{S}_a\,(\mathbf{I}-\mathbf{A})^T, \tag{5.109}$$

where $\mathbf{S}_a$ is the *a priori* covariance matrix. Ideally, $\mathbf{S}_a$ is calculated from an ensemble of states that also include the retrieved state (Rodgers and Connor, 2003). However, for the application of aerosol satellite remote sensing $\mathbf{S}_a$ is in general not known, which makes it difficult to calculate $\mathbf{S}_r$ for individual retrievals. However, an estimate for the upper boundary of the regularization error can be obtained by calculating $\mathbf{S}_r$ from Eq. (5.109) by assuming an *a priori* covariance matrix representing maximum values for the errors on the elements of $\mathbf{x}_a$. In order to estimate the maximum errors on the elements of $\mathbf{x}_a$ we used the 17 tropospheric aerosol models of Torres et al. (2001). For these 17 aerosol models we calculated the mean value and considered the maximum difference between the mean value and the actual value as *a priori* error. This resulted in the following *a priori* standard deviations σ_a for the different parameters: $\sigma_a(r^s_{\text{eff}}) = 0.05\ \mu\text{m}, \sigma_a(v^s_{\text{eff}}) = 0.23, \sigma_a(r^l_{\text{eff}}) = 1.29\ \mu\text{m}, \sigma_a(v^l_{\text{eff}}) = 0.22, \sigma_a(m_r) = 0.065$, and $\sigma_a(m_i) = 0.01$. For z_b and C_{pig} we assumed *a priori* errors of 100%.

The contribution matrix $\mathbf{D}$ plays an important role for calculating the error propagation from measurement $\mathbf{y}$ to state vector $\mathbf{x}$. Assuming that the forward model is linear within the range of the errors, the effect of a random measurement error on the state vector is called retrieval noise. The retrieval noise covariance matrix $\mathbf{S}_x$ is given by

$$\mathbf{S}_x = \mathbf{D}\ \mathbf{S}_y\ \mathbf{D}^T. \tag{5.110}$$

Systematic state vector errors $\Delta\mathbf{x}$ due to systematic measurement errors $\Delta\mathbf{y}$ can also be evaluated using the contribution matrix:

$$\Delta\mathbf{x} = \mathbf{D}\ \Delta\mathbf{y}, \tag{5.111}$$

and a similar expression holds for systematic forward model errors $\Delta\mathbf{F}$, but with $\Delta\mathbf{y}$ replaced by $-\Delta\mathbf{F}$. Of course, the systematic errors in measurement and forward model are not known, because otherwise they would have been corrected for. However, examples of systematic state vector errors can be calculated for some reasonable scenarios of systematic measurement and forward model errors.

For estimating direct radiative forcing by aerosols, aerosol optical properties such as (spectral) optical thickness and single scattering albedo are very important. These optical properties can be derived from the microphysical aerosol parameters contained in the state vector $\mathbf{x}$. The standard deviation σ_τ on the optical thickness can be obtained from the retrieval noise covariance matrix $\mathbf{S}_x$ via

$$\sigma_\tau = \sqrt{\sum_{i=1}^{N}\sum_{j=1}^{N}\ s_{i,j}\ \frac{\partial\tau}{\partial x_i}\ \frac{\partial\tau}{\partial x_j}}, \tag{5.112}$$

where $s_{i,j}$ denotes element (i,j) of $\mathbf{S}_x$. The effect of the regularization error covariance matrix can be obtained in the same way. Systematic errors $\Delta\tau$ on the aerosol optical thickness τ are given by

$$\Delta\tau = \sum_{i=1}^{N} \Delta x_i\ \frac{\partial\tau}{\partial x_i}. \tag{5.113}$$

Expressions similar to Eqs (5.112) and (5.113) hold for the aerosol single scattering albedo ω.

For completeness we would like to note that if in Eq. (5.101) the *a priori* state vector is included in the side constraint with $\gamma = 1$, and the inverse of the *a priori* covariance matrix is used for $\mathbf{\Gamma}$, Eq. (5.101) would be identical to the optimal estimation solution (Rodgers, 1976). The optimal estimation solution represents the maximum likelihood solution given the measurement, *a priori* information, and the corresponding covariance matrices. As stated above for the application of satellite aerosol remote sensing the *a priori* covariance matrix is not known with useful accuracy. If an *ad hoc* matrix is used that has been constructed as a rough estimate of the *a priori* covariance matrix, the optimal estimation solution reduces to the Phillips–Tikhonov solution, but with an arbitrarily chosen value for the regularization parameter (i.e. $\gamma = 1$) which is not necessarily close to the corner of the L-curve. In the Phillips–Tikhonov minimization (5.101), the matrix $\mathbf{\Gamma}$ is also an *ad hoc* matrix, but its absolute values do not affect the solution, because the side constraint is weighted by the regularization parameter, for which a suitable value is found using the L-curve. Therefore, in this chapter we prefer the use of Phillips–Tikhonov regularization in combination with the L-curve criterion for choosing the regularization parameter, instead of the use of the optimal estimation method with an *ad hoc* matrix as a rough estimate of the *a priori* covariance matrix.

5.6.2 Levenberg–Marquardt iteration

In general, the inversion of Eq. (5.1) represents a highly non-linear problem. Therefore, if the first guess state vector $\mathbf{x}_o$ is too far from the true state vector the linear approximation in Eq. (5.99) may be poor. In that case, the inversion of the linearized forward model in Eq. (5.99) may result in a new state vector that yields a higher χ^2 difference between forward model $\mathbf{F}(\mathbf{x}_{n+1})$ and measurement $\mathbf{y}$ than the first guess state vector $\mathbf{x}_o$, i.e. a step has been taken in the wrong direction. In order to prevent the inversion from taking a large step away from the minimum χ^2, we use the Levenberg–Marquardt method (Levenberg, 1944; Marquardt, 1963), which minimizes the step size between two iteration steps in addition to minimizing the difference between linearized forward model and measurement. Thus, for the first few iteration steps we replace our original cost function (5.101) by the Levenberg–Marquardt cost function, given by

$$\hat{\mathbf{x}}_{n+1}(\nu) = \min_{\mathbf{x}} \left(||(\mathbf{K}\mathbf{x} - \tilde{\mathbf{y}})||^2 + \nu \, ||\mathbf{\Gamma}(\mathbf{x} - \mathbf{x}_n)||^2 \right), \tag{5.114}$$

where the subscripts n and $n+1$ denote the current and next iteration step, respectively, ν is a parameter that controls the step size, and $\mathbf{\Gamma}$ is the same weighting factor matrix as in Eq. (5.101). Ideally, ν should be chosen such that the step taken yields an optimal improvement in χ^2 between $\mathbf{F}(\mathbf{x}_{n+1})$ and $\mathbf{y}$. However, in order to find that value of ν, it is necessary to evaluate χ^2 for a large number of trial values of ν, which requires a lot of computation time because for each trial ν a new forward model calculation is required. A more efficient application of Eq. (5.114) was proposed by Press et al. (1992). They start with a certain first guess value for ν and evaluate for that ν the χ^2 between

$\mathbf{F}(\mathbf{x}_{n+1})$ and $\mathbf{y}$. If this χ^2 is smaller than the χ^2 of the previous iteration step, they proceed the iteration with $\mathbf{x}_{n+1}$ and decrease ν by a certain factor. If χ^2 is larger than the χ^2 of the previous iteration step, they proceed the iteration with $\mathbf{x}_n$ and increase ν by a certain factor. For our aerosol retrieval problem also this more efficient method still requires a large number of iteration steps (often more than 40 iteration steps are needed). Therefore, we developed a method to speed up the iteration process considerably. Here, for each iteration step we solve Eq. (5.114) for many different (say 50) values of ν. Each ν results in a different retrieved state vector $\mathbf{x}_{n+1}$. For all these retrieved state vectors we use an approximate forward model to evaluate (an approximation of) the χ^2 difference between forward model $\mathbf{F}(\mathbf{x}_{n+1})$ and measurement $\mathbf{y}$. This approximate forward model makes use of the fact that the forward model is more linear as a function of the discretized size distribution than as a function of r_{eff} and v_{eff}, viz.

$$\mathbf{F}_{\text{appr}}(\mathbf{x}) = \mathbf{F}(\mathbf{x}_n) + \sum_{j=1}^{2} \sum_{i=1}^{N} \frac{\partial \mathbf{F}}{\partial n_i} \Delta n_i + \frac{\partial \mathbf{F}}{\partial \mathbf{x}_{\text{rest}}} \Delta \mathbf{x}_{\text{rest}}. \tag{5.115}$$

Here, n_i is the value of the size distribution in size bin i, Δn_i is a change in n_i caused by a change in r_{eff} and v_{eff}, and $\mathbf{x}_{\text{rest}}$ is the part of the state vector not including r_{eff} and v_{eff}. Furthermore, the summation over j represents the summation over the two modes of the size distribution and the summation over i from 1 to N represents the summation over all size bins. The forward model $\mathbf{F}_{\text{appr}}$ of Eq. (5.115) is a better approximation of the true forward model $\mathbf{F}$ than the linear approximation of Eq. (5.2), because the Δn_i in Eq. (5.115) are calculated using the non-linear expression for the size distribution $n_i(r_{\text{eff}}, v_{\text{eff}})$. Using, the approximate forward model (5.115) we efficiently estimate the value ν_{opt} which yields optimal improvement in χ^2, and use the corresponding state vector $\mathbf{x}_{n+1}(\nu_{\text{opt}})$ to proceed the iteration.

We follow the iteration process described above until χ^2 becomes smaller than a certain threshold. For values of χ^2 below this threshold we assume that the problem has become sufficiently linear for us to be able to replace Eq. (5.114) by the original cost function Eq. (5.101). In the final iteration step, this minimization yields the final solution of our retrieval problem with the corresponding regularization and error analysis.

5.7 Application to GOME-2

In this section we apply the retrieval concept presented in the previous subsections to synthetic measurements of intensity and polarization of the Polarization Measuring Device (PMD) of GOME-2. The first of three GOME-2 satellite instruments has been launched in October 2006 on the Eumetsat's Metop satellite. In total, a GOME-2 measurement series will be performed till 2020. Here, we will first demonstrate that the proposed retrieval concept is well suited to solve the non-linear aerosol retrieval problem. Furthermore, we will demonstrate that a linear error mapping procedure, as used in Eq. (5.110), allows a sound error analysis for the application of aerosol retrieval. Finally, we will present an analysis of the information content of GOME-2 measurements of intensity and polarization, including a comparison with the information content of intensity -only retrievals from the GOME-2 PMD.

5.7.1 GOME-2 measurements

The PMD of GOME-2 measures the 312–800 nm spectral range using 200 detector pixels with a spectral resolution of 2.8 nm at 312 nm and about 40 nm at 800 nm. The components of the intensity polarized parallel and perpendicular to the optical plane are measured simultaneously. These components are denoted by I_i^l and I_i^r, respectively, for detector pixel i. The design of the PMD was driven by the optical identity of the l- and r-channels. Given the need of a lightweight instrument, a prism spectrometer was chosen instead of a more complex grating solution. The measurement I_i^l is simulated by

$$I_i^l = \int_0^\infty \mathrm{d}\lambda \; S_i(\lambda) \; I^l(\lambda), \tag{5.116}$$

where the integration over wavelength λ describes the effect of a Gaussian spectral response function $S(\lambda)$, where $I^l(\lambda)$ denotes the l-component of the intensity at the entrance of the instrument. The measurement I_i^r is simulated in the same manner as in Eq. (5.116). From the measured intensities I_i^l and I_i^r the Stokes parameters I_i and Q_i can be obtained, viz.,

$$I_i = I_i^l + I_i^r, \tag{5.117}$$
$$Q_i = I_i^l - I_i^r. \tag{5.118}$$

The simulated values of I_i and Q_i are superimposed by a random Gaussian noise. Here, we calculated the contributions of photon-shot-noise and detector-noise using the transmission properties of GOME-2, known from the on-ground calibration. However, we believe that considering only these two error sources an unrealistically positive retrieval diagnose will be obtained, because the measurement possibly also contains other errors which may introduce a random-like structure, such as errors due to spatial aliasing, unknown spectral features introduced by the diffuser plate, and spectral calibration errors. Also the forward model may contain random-like errors, for example due to errors in accounting for molecular absorption, the description of underlight, the assumed distribution of surface slopes of the oceanic waves, and the prediction of whitecap coverage from wind speed. Therefore, in addition to the contributions of photon-shot-noise and detector-noise, a noise floor of 1% is added to the simulations of I_i and Q_i, to account for such errors.

Due to limitations in the GOME-2 data rate the information of the 200 detector pixels has to be co-added onboard to form 15 programmable bands. The expected wavelength ranges of these bands are denoted in Table 5.1. For

Table 5.1. Probable PMD band selection for GOME-2 (Hasekamp et al., 2004a). The wavelength range is indicated by the center wavelengths of the first and last PMD pixel, respectively, not including the Full Width at Half Maximum (FWHM) of the slit function, which is shown in the last column

Band no.	Wavelength range (nm)	No. of pixels	FWHM (nm)
1	spare	–	–
2	311.8–314.3	5	3.1
3	316.9–318.8	4	3.3
4	321.5–329.3	12	3.5
5	330.8–334.6	6	3.8
6	336.2–340.0	24	4.8
7	361.0–377.9	20	4.8
8	380.3–383.9	4	6.1
9	399.8–428.0	19	7.8
10	435.1–493.5	23	10.2
11	495.5–552.4	23	12.5
12	567.9–598.0	5	25.2
13	600.0–660.0	11	30.0
14	743.1–766.6	3	38.5
15	783.6–792.4	2	43.9

aerosol retrieval it is anticipated that band 6–15 will be used. The intensity I_{pmd} for a given PMD band is given by

$$I_{\mathrm{pmd}} = \sum_{i=1}^{N} I_i, \tag{5.119}$$

where the summation in Eq. (5.119) describes the co-adding over a number of N detector pixels. The Stokes parameter Q_{pmd} is obtained in the same manner. The standard deviation σ_{pmd} for the Gaussian error on I_{pmd} or Q_{pmd} is given by

$$(\sigma_{\mathrm{pmd}})^2 = \frac{1}{N}\left(\sum_{i=1}^{N} (\sigma_i)^2\right), \tag{5.120}$$

where σ_i is the standard deviation for a detector pixel.

In this study we consider retrievals using intensity as well as polarization measurements and retrievals using only intensity measurements. For the latter retrievals we also use the PMD spectral bands of Table 5.1. In this comparison we want to avoid differences in information content due to the fact that adding an extra set of intensity measurements improves the signal-to-noise ratio by a factor $\sqrt{2}$. Therefore, for the calculation of the measurement error covariance matrix for the intensity -only retrievals we used the signal-to-noise ratio corresponding to a double set of intensity measurements.

5.7.2 Retrieval results

The retrieval procedure described in this chapter is tested on a set of 100 synthetic GOME-2 PMD measurements of intensity and polarization, created for randomly chosen aerosol parameters within a specified range. The chosen ranges for the different parameters were: 0.1–0.2 μm for r_{eff} of the small mode, 0.65–3.40 μm for r_{eff} of the large mode, 0.16–0.65 for v_{eff} of the small mode, 0.5–0.9 for v_{eff} of the large mode, 1.4–1.6 for m_r, $5 \cdot 10^{-7}$–0.02 for $|m_i|$, and aerosol columns for both modes corresponding to an optical thickness at 550 nm in the range 0.05–0.5. The oceanic pigment concentration ranged from 0.5–2 mg/m^3, whereas the height z_b of the layer where the bulk of the aerosols is located (see Eq. 5.98) was kept fixed at 2 km.

In all cases the iteration converged to a stable solution and also the χ^2 difference between forward model and measurement was close to 1 in all cases. Figure 5.4 shows the retrieved optical thickness (i.e. derived from the retrieved parameters) versus the true optical thickness at 550 nm. It can be seen that the retrieved optical thickness corresponds well to the true optical thickness. In general, the agreement is within 5%. This example indicates that the implemented retrieval approach is suited for aerosol retrieval.

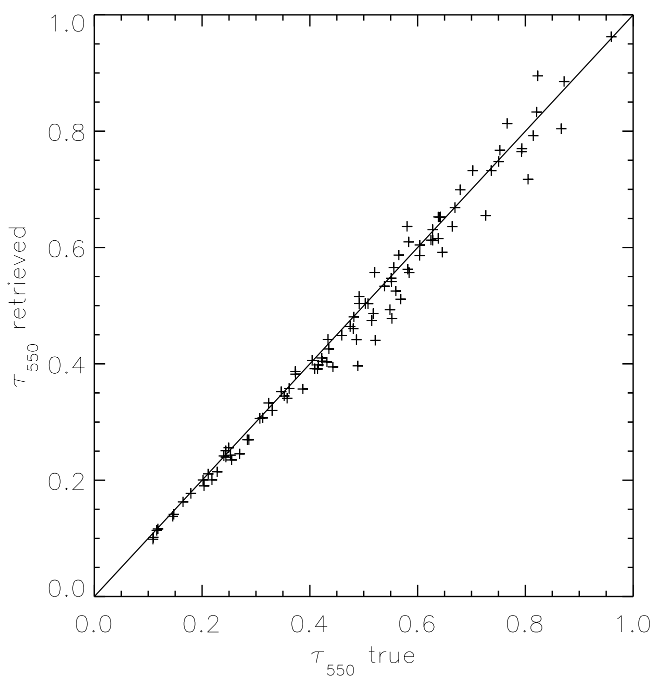

Fig. 5.4. Retrieved optical thickness versus true optical thickness at 550 nm for 100 synthetic retrievals.

An error analysis based on a linearized forward model, as is performed in Eq. (5.110), is only valid if the forward model is (approximately) linear within the error range. In order to test if a sound error analysis can be performed using a linearized forward model, we investigate whether the differences between the retrieved state vector and the true state vector are consistent with the retrieval noise covariance matrix calculated by Eq. (5.110). In order to exclude the effect of the *a priori* state vector from the comparison we replace $\mathbf{x}_a$ by $\mathbf{x}_{\text{true}}$ in the term $(\mathbf{I}-\mathbf{A})\mathbf{x}_a$ in Eq. (5.107) for the final iteration step. If the differences between the retrieved state vector and the true state vector are consistent with the retrieval noise covariance matrix (5.110), then the distribution of $(x_i^r - x_i^t)/\sigma_i$, where x_i^r is the ith element of the retrieved state vector, x_i^t is the corresponding true value, and σ_i the standard deviation that follows from Eq. (5.110), is given by the standard Gaussian distribution

$$f(y) = \frac{\exp(-y^2/2)}{\sqrt{2\pi}}. \tag{5.121}$$

For the retrievals on the 100 synthetic measurements described above, the corresponding distribution is shown in Fig. 5.5. The distribution shown contains the values $(x_i^r - x_i^t)/\sigma_i$ for all state vector elements. From Fig. 5.5 it follows that the distribution of the retrieved aerosol parameters reproduces the standard Gaussian distribution well. This demonstrates that the linear approximation of Eq. (5.110) is valid for the calculation of the retrieval noise covariance matrix. So, a linear error mapping procedure can be used to investigate the retrieval capabilities of a given instrument concept, without doing a full iterative retrieval. The linearized radiative transfer model described in this chapter is a powerful tool for this purpose.

5.7.3 Information content

We investigated the information content of GOME-2 measurements using linearized radiative transfer calculations for the two aerosol types in Table 5.2, where the optical thickness at 550 nm $\tau_{550} = 0.3$. Here, aerosol type A corresponds to biomass-burning aerosols and type B corresponds to oceanic aerosols. Figure 5.6 shows the DFS as a function of viewing zenith angle (VZA) for a solar zenith angle (SZA) of 40° and a relative azimuth angle of 60° (positive VZA) or −120° (negative VZA), for retrievals using intensity as well as polarization measurements (left panel) and for retrievals using only intensity measurements (right panel). From this figure it follows that for retrievals using intensity as well as polarization measurements the DFS is in the range 6–8 which is 1–4 degrees higher than for retrievals using only intensity measurements. So, the use of polarization measurements significantly improves the information content.

In order to interpret by which parameters the DFS is mainly determined, we show in Fig. 5.7 the derivatives of the retrieved parameters with respect to their *a priori* values for the biomass-burning aerosol type. Here, if for a parameter $\partial x/\partial x_a = 1$, this parameter is fully determined by its *a priori* value, whereas if $\partial x/\partial x_a = 0$ the parameter is not influenced by its *a priori* value at all. Here, it

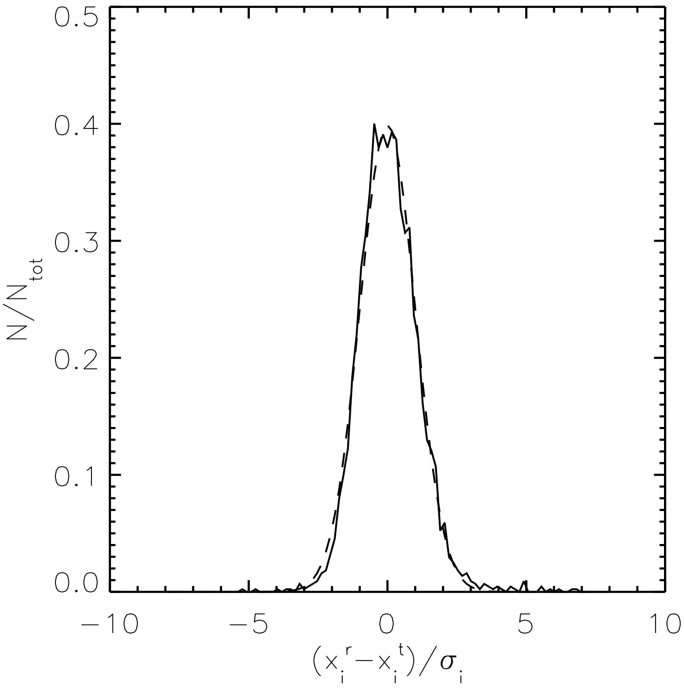

Fig. 5.5. Distribution of $(x_i^r - x_i^t)/\sigma_i$ for the 100 retrievals on synthetic GOME-2 measurements (solid line). N/N_{tot} indicates the number of points in a certain size bin normalized to the total number of points. The standard Gaussian distribution of Eq. (5.121) is given by the dashed line. The distribution contains 101 bins between -8 and 8.

Table 5.2. Aerosol types used to create synthetic measurements of intensity and polarization. The aerosol types are adopted from Torres et al. (2001). Type A corresponds to biomass-burning aerosols and type B corresponds to oceanic aerosols. See Appendix B for definitions of r_{eff} and v_{eff}. f_l denotes the fraction of large mode particles. Concerning the aerosol altitude distribution of Eq. (5.98), $z_b = 2$ km and $z_t = 10$ km

Type	r^s_{eff}	v^s_{eff}	r^l_{eff}	v^l_{eff}	f_l	$\tau^l_{550}/\tau^{\mathrm{tot}}_{550}$	m_r	m_i	τ_{350}	τ_{550}	ω_{350}
A	0.119	0.174	2.671	0.704	$2.05 \cdot 10^{-4}$	0.087	1.50	-0.02	0.657	0.300	0.892
B	0.105	0.651	0.840	0.651	$1.53 \cdot 10^{-2}$	0.851	1.40	$-5 \cdot 10^{-8}$	0.336	0.300	1.000

is important to note that due to our choice of the matrix $\mathbf{\Gamma}$ in Eq. (5.101) we force the aerosol columns of both modes to be fully independent of their *a priori* values, i.e. $\partial x/\partial x_a = 0$ for these parameters. Therefore, these parameters are not included in Fig. 5.7. It follows from Fig. 5.7 that the polarization measurements mostly add information on the effective variance of the small mode, the imagi-

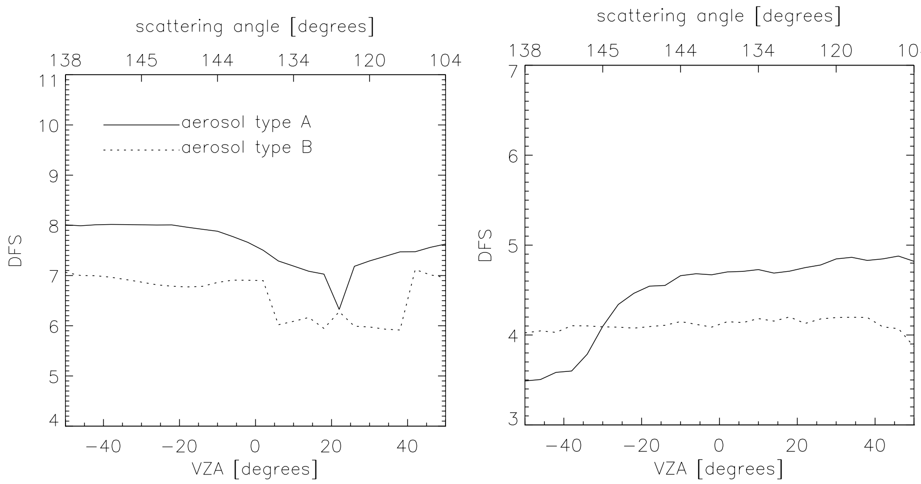

Fig. 5.6. Degrees of Freedom for Signal (DFS) for retrievals from GOME-2 PMD intensity and polarization measurements (left panel) and for retrievals using only intensity measurements (right panel). The two aerosol types of Table 5.2 are used. The DFS is shown as a function of viewing zenith angle (VZA) for a solar zenith angle of 40^o and a relative azimuth angle $\varphi_o - \varphi_v = 60^o$ for positive VZA and $\varphi_o - \varphi_v = -120^o$ for negative VZA. An oceanic pigment concentration of 1 $\mathrm{mg/m^3}$ was used for the simulations.

nary part of the refractive index, and the height z_b of the layer where the bulk of the aerosols is located. Furthermore, significant additional information can be retrieved on the effective radius of the small mode and the oceanic pigment concentration. Both retrievals contain little information on the size distribution parameters of the large mode. This can be explained by the fact that for the biomass-burning aerosol type the contribution of the large mode to the total optical thickness is relatively small. In contrast, the oceanic aerosol type (not shown) is dominated by the large mode. Therefore, for this aerosol type more information is available on the effective radius of the large mode, while the effective radius of small mode depends stronger on *a priori* information. However, the effective variance of both modes strongly depends on *a priori* for the oceanic aerosol type.

From Fig. 5.7 we conclude that the use of polarization measurements makes it possible to retrieve information on aerosol size and refractive index that cannot be retrieved using only intensity measurements. This can be explained by the characteristic sensitivity of polarization properties of light to aerosol microphysical properties, as shown, for example, by Hansen and Travis (1974). Furthermore, as follows from Fig. 5.7, polarization measurements allow the retrieval of information on aerosol height. This information mainly comes from measurements at wavelengths below about 450 nm, where the Rayleigh scattering optical thickness is relatively large. Since most Rayleigh scattering takes place low in the atmosphere, the Rayleigh scattering signal is more strongly attenuated if

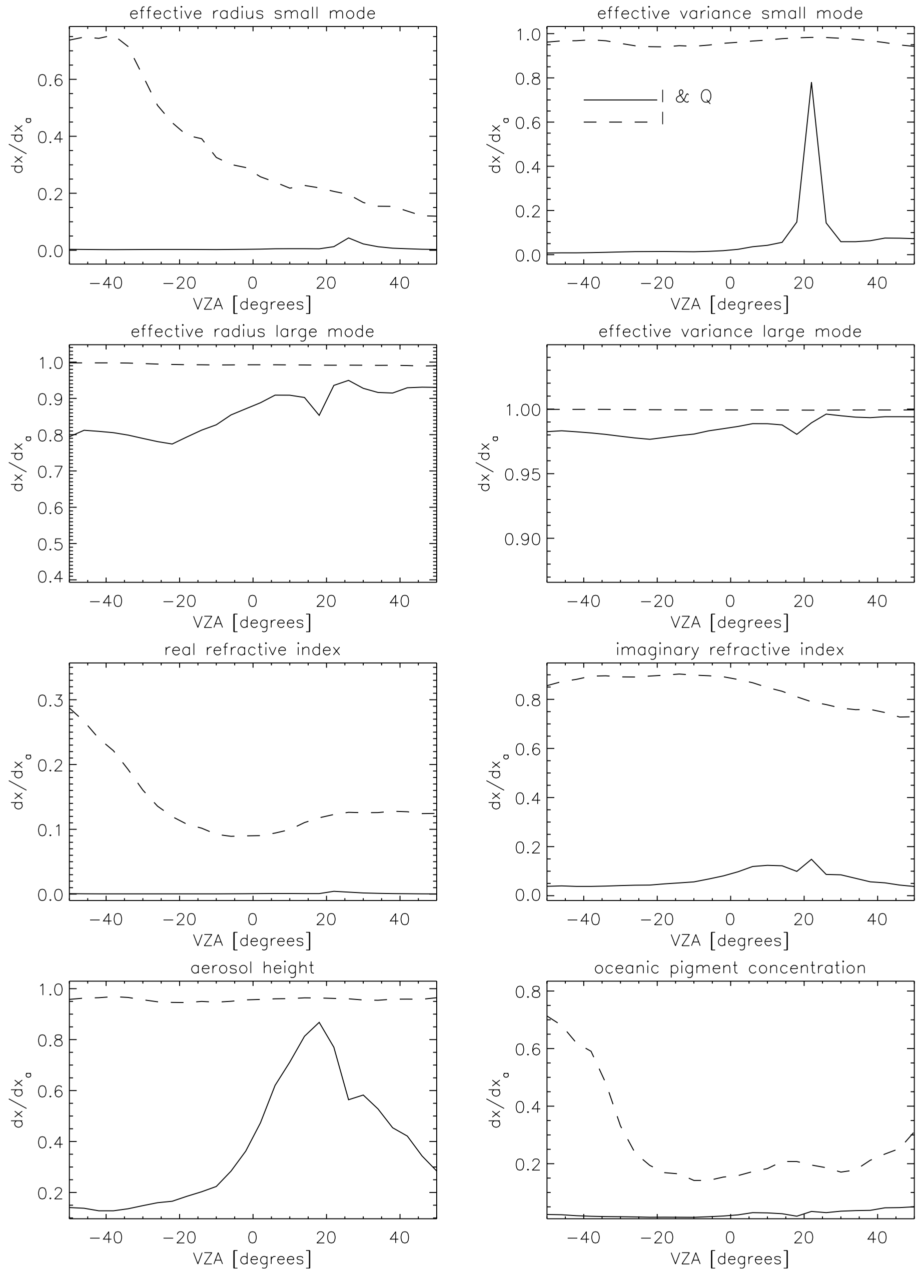

Fig. 5.7. Derivatives of retrieved values with respect to their *a priori* values for the biomass-burning aerosol type (A), as a function of viewing zenith angle. Other angles as in Fig. 5.6.

the aerosols are located higher in the atmosphere. So, the degree of polarization of the backscattered light becomes lower for increasing aerosol height, because Rayleigh scattering generally causes a higher degree of polarization than aerosol scattering. For the oceanic aerosol type, there is significantly less information available on aerosol height (not shown) then for the biomass-burning aerosol type, because for the oceanic aerosol type the aerosol optical thickness at short wavelengths is much smaller than for biomass-burning aerosols. In addition to the aerosol parameters, the oceanic pigment concentration also can be retrieved using GOME-2 intensity and polarization measurements. This is due to the characteristic spectral signature of oceanic pigment.

Figure 5.8 shows for aerosol type A the total retrieval error (retrieval noise and regularization error) on the optical thickness at 350 nm and 550 nm, respectively, for retrievals using intensity and polarization measurements and retrievals using only intensity measurements. It can be seen that for the retrievals using intensity and polarization measurements, the optical thickness error shows a distinct maximum around a VZA of 20°. The reason for this is that at these geometries the sensitivity of Stokes parameter Q to atmospheric properties is rather low, which means that here the retrievals rely for a large part on intensity measurements. This strong dependence on viewing geometry demonstrates that the aerosol information retrieved using single-viewing-angle polarization measurements is for some geometries less useful than for other geometries. These geometries are well defined and the corresponding aerosol retrieval products should be labeled as less reliable. Away from this maximum, the optical thickness error is around 0.025 (3.7%) at 350 nm and around 0.017 (5.7%) at 550 nm. The optical thickness errors for retrievals using only intensity measurements are

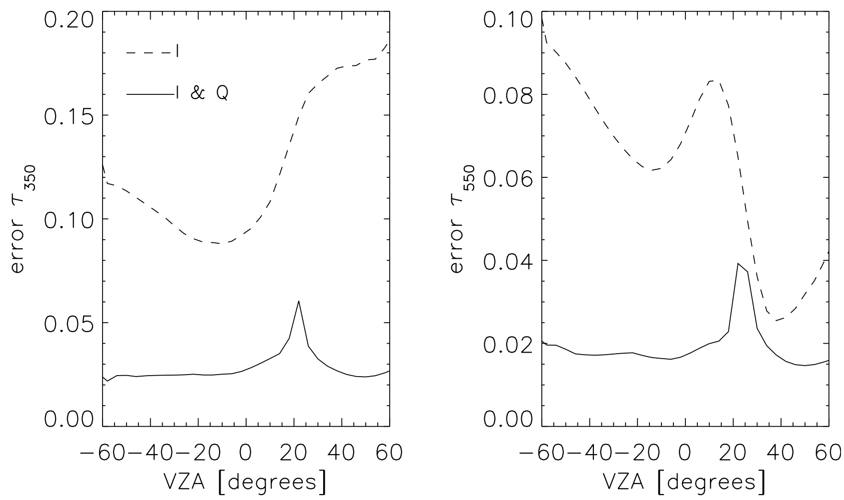

Fig. 5.8. Total retrieval error on the retrieved optical thickness at 350 nm (left panel) and 550 nm (right panel), as a function of viewing zenith angle. Other angles as in Fig. 5.6.

a factor 2–7 higher. The increase in total optical thickness error mainly comes from the regularization error, which means that optical thickness retrievals from only intensity measurements are very sensitive to *a priori* information on aerosol size distribution and refractive index. So, the additional information on aerosol size distribution and refractive index (see Fig. 5.7) that can be retrieved including polarization measurements, is not only important information on its own, but is also essential if a reliable optical thickness retrieval is to be obtained.

Figure 5.9 shows the total retrieval error on the aerosol single scattering albedo ω_{350} at 350 nm, for retrievals using intensity as well as polarization measurements, and retrievals using only intensity measurements. Here, we use a wavelength of 350 nm because information about aerosol single scattering albedo mainly comes from shorter wavelengths due to interaction with Rayleigh scattering (Torres et al., 1998). For retrievals using intensity as well as polarization measurements, the total error on the single scattering albedo is mostly below 0.015. For intensity -only retrievals the total error on ω_{350} is about a factor 2–4 larger than for retrievals using polarization measurements. So, the retrieval of single scattering albedo also benefits significantly from including polarization measurements.

To summarize, multi-wavelength single-viewing-angle measurements of intensity as well as polarization in the range 340–800 nm contain valuable information on aerosol size, refractive index, spectral optical thickness, and UV-single scattering albedo. These aerosol characteristics are of essential importance for climate research. Using only intensity measurements in the same spectral range, significantly less information on microphysical aerosol properties can be retrieved,

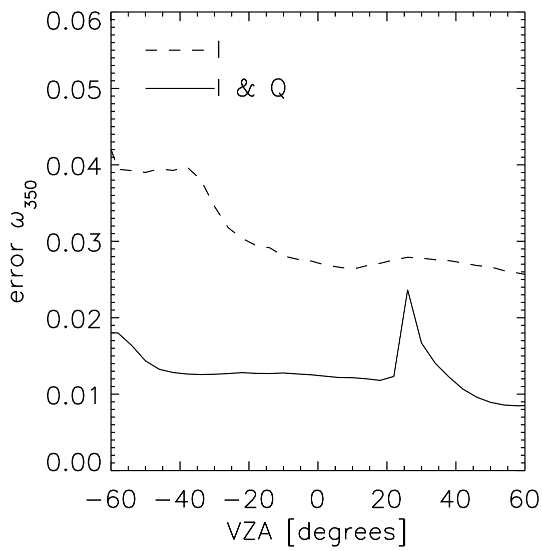

Fig. 5.9. Total retrieval error on the retrieved single scattering albedo at 350 nm as a function of viewing zenith angle. Other angles as in Fig. 5.6.

leading to (much) larger errors on the corresponding retrieved optical thickness and single scattering albedo. These conclusions have also been tested for other aerosol types, and similar results were obtained. In addition, multi-viewing-angle measurements will provide information on the aerosol phase matrix which in turn will provide additional constraints on microphysical aerosol properties and on surface reflectance properties.

5.8 Conclusions

The analytical linearization of vector radiative transfer with respect to physical aerosol properties and its use in satellite remote sensing have been reviewed. The linearization consists of two steps. The first step is the calculation of the derivatives of the four Stokes parameters at the top of the atmosphere with respect to scattering coefficient, absorption coefficient, and the expansion coefficients of the scattering phase matrix. These derivatives are calculated analytically employing the forward-adjoint perturbation theory. Here, general expressions are presented that can be applied for the linearization of any vector radiative transfer model that calculates the internal radiation field in the model atmosphere. The second step is the calculation of the derivatives of the scattering coefficient, absorption coefficient, and the expansion coefficients of the scattering phase matrix, with respect to the real and imaginary part of the refractive index, and parameters describing the size distribution (e.g. effective radius, effective variance). These derivatives are analytically calculated following Mie theory. The developed linearization approach has been implemented in a Gauss–Seidel vector radiative transfer model. The linearized radiative transfer model has been incorporated in a retrieval algorithm based on the Phillips–Tikhonov regularization method in combination with the Levenberg–Marquardt iterative method. This retrieval algorithm aims to retrieve microphysical aerosol parameters corresponding to a bi-modal aerosol size distribution. Additionally, the oceanic pigment concentration and information on aerosol height are retrieved from the measurement. We used synthetic GOME-2 measurements of intensity and polarization to demonstrate that the developed iterative retrieval approach based on linearized radiative transfer is well suited to solve the non-linear aerosol retrieval problem. Furthermore, we demonstrated that a linear error mapping procedure can be used to perform a solid error analysis, without doing a full iterative retrieval.

Finally, we presented and overview of the information content of GOME-2 measurements of intensity and polarization. Here, we considered the retrieval of nine aerosol parameters corresponding to a bi-modal aerosol size distribution: the column integrated aerosol number concentration of both modes, the effective radius of both modes, the effective variance of both modes, the real- and imaginary part of the refractive index, and the height of the layer where the bulk of the aerosols is located. In addition to the nine aerosol parameters we also considered the oceanic pigment concentration as an unknown parameter. It is demonstrated that for this retrieval setup the DFS is in the range 6–8. Here, the aerosol loading of both modes, the effective radius of at least one mode, the real and imaginary part of the refractive index, the height of the layer where the bulk of the aerosols is located, and the oceanic pigment concentration can for most viewing geometries be retrieved from the measurement with negligible dependence on *a priori* information.

For retrievals that only use intensity measurements the DFS is significantly less than for retrievals also using polarization measurements, namely in the range 3.5–5. For these retrievals no significant information on aerosol imaginary refractive index, effective variance, and aerosol height can be retrieved. Furthermore, the information on effective radius, real part of the refractive index, and oceanic pigment concentration is much more affected by *a priori* information than retrievals that include polarization measurements.

To conclude, the results of this chapter demonstrate that a linearized radiative transfer model as presented here provides a powerful tool for efficiently solving the aerosol retrieval problem, and additionally for a solid error analysis. Using this tool, we showed that multi-wavelength single-viewing-angle measurements of intensity as well as polarization in the range 340–800 nm contain valuable information on aerosol size, refractive index, spectral optical thickness, and UV-single scattering albedo. Using only intensity measurements in the same spectral range significantly less information on microphysical aerosol properties can be retrieved, leading to (much) larger errors in the corresponding retrieved optical thickness and single scattering albedo. The retrievals can be further improved using multiple-viewing-angle measurements and highly spectrally resolved measurements in absorption bands of well mixed atmospheric gases, such as oxygen.

Appendix A: The Mie coefficients and their derivatives

The Mie coefficients a_n and b_n are calculated using the method of de Rooij and van der Stap (1984). Here, we will summarize the relevant formulas and for further details we refer to the corresponding paper. Furthermore, we give expressions for the derivatives of a_n and b_n with respect to the real and imaginary part of the refractive index, used in section 5.4.2. The Mie coefficients are given by (see, for example, Deirmendjian (1969)):

$$a_n = \frac{(D_n(z)/m + n/x)\ \Psi_n(x) - \Psi_{n-1}(x)}{(D_n(z)/m + n/x)\ \zeta_n(x) - \zeta_{n-1}(x)}, \tag{5.122}$$

$$b_n = \frac{(mD_n(z) + n/x)\ \Psi_n(x) - \Psi_{n-1}(x)}{(mD_n(z) + n/x)\ \zeta_n(x) - \zeta_{n-1}(x)}, \tag{5.123}$$

where $m = m_r + im_i$ is the complex refractive index, x is the size parameter $2\pi r/\lambda$, and $z = mx$. Furthermore,

$$\Psi_n(x) = xj_n(x), \tag{5.124}$$

$$\zeta_n(x) = \Psi_n(x) + i\chi_n(x), \tag{5.125}$$

with

$$\chi_n(x) = -xy_n(x), \tag{5.126}$$

where $j_n(x)$ and $y_n(x)$ are the spherical Bessel functions of the first and second kind, respectively. $D_n(z)$ is the only function that depends on refractive index and is given by

$$D_n(z) = \frac{d}{dz} \ln \Psi_n(z) = -\frac{n}{z} \frac{\Psi_{n-1}(z)}{\Psi_n(z)}. \tag{5.127}$$

The functions $\Psi_n(x)$ and $\chi_n(x)$ and $D_n(z)$ are all calculated using recurrence relations. Here, $\chi_n(x)$ is calculated by upward recursion using the recurrence relation

$$\chi_{n+1}(x) = \frac{2n+1}{x} \chi_n(x) - \chi_{n-1}(x), \tag{5.128}$$

with initial functions

$$\chi_{-1}(x) = \sin x, \qquad \chi_0(x) = \cos x. \tag{5.129}$$

$\Psi_n(x)$ is calculated using downward recursion:

$$\Psi_n(x) = r_n(x)\, \Psi_{n-1}(x), \tag{5.130}$$

where

$$r_n(x) = \left[\frac{2n+1}{x} - r_{n+1}(x)\right]^{-1}. \tag{5.131}$$

The recursion is started at $n = N_1(x)$ where

$$N_1(x) = x + 4.05x^{1/3} + 60, \tag{5.132}$$

and $r_{N_1}(x) = 0$ (de Rooij and van der Stap, 1984).

$D_n(z)$ is calculated using the following downward recursion relation:

$$D_n(z) = \frac{n+1}{z} - \left(D_{n+1}(z) + \frac{n+1}{z}\right)^{-1}, \tag{5.133}$$

where the recursion is started at $n = N_2(z)$ with

$$N_2(z) = z + 4.05z^{1/3} + 10, \tag{5.134}$$

and $D_{N_2}(z) = 0$ (de Rooij and van der Stap, 1984).

The derivatives of a_n and b_n with respect to the real and imaginary part of the the refractive index are given by

$$[a_n]' = \frac{\left([D_n(z)]'/m - D_n(z)/m^2\right)(\Psi_{n-1}\zeta_n - \Psi_n\zeta_{n-1})}{[(D_n(z)/m + n/x)\ \zeta_n(x) - \zeta_{n-1}(x)]^2}, \tag{5.135}$$

$$[b_n]' = \frac{m[D_n(z)]'\,(\Psi_{n-1}\zeta_n - \Psi_n\zeta_{n-1})}{[(mD_n(z) + n/x)\ \zeta_n(x) - \zeta_{n-1}(x)]^2}, \tag{5.136}$$

where the prime indicates the derivative with respect to either m_r or im_i. Here, it is important to note that in section 5.4.2 we use the derivatives with respect to m_i which follow directly from the here given derivatives with respect to im_i. The derivative $[D_n(z)]'$ is found by backward recursion via

$$[D_n(z)]' = \frac{-x(n+1)}{z^2} + \left([D_{n+1}(z)]' + \frac{-x(n+1)}{z^2}\right)\left(D_{n+1}(z) + \frac{n+1}{z}\right)^{-2}, \tag{5.137}$$

starting the recursion at N_2 with $[D_{N_2}(z)]' = 0$.

Appendix B: Aerosol and ocean properties

B.1 Aerosol size distribution

For all simulations in this chapter we assume that the aerosol size distribution is bi-modal, where the size distribution n for each mode is given by a lognormal function

$$n(r) = \frac{1}{\sqrt{2\pi}\,\sigma_g\, r} \exp\left[-(\ln r - \ln r_g)^2/(2\sigma_g^2)\right], \tag{5.138}$$

where r describes particle radius,

$$\ln r_g = \int_0^\infty \ln r\; n(r)\, \mathrm{d}r, \tag{5.139}$$

and

$$\sigma_g^2 = \int_0^\infty (\ln r - \ln r_g)^2\; n(r)\, \mathrm{d}r. \tag{5.140}$$

As shown by Hansen and Travis (1974) it is useful to characterize (a mode of) the size distribution by the effective radius r_{eff} and effective variance v_{eff}, because these parameters are relatively independent from the actual shape of the distribution. Here,

$$r_{\mathrm{eff}} = \frac{1}{G} \int_0^\infty r\pi r^2 n(r)\, \mathrm{d}r, \tag{5.141}$$

and

$$v_{\mathrm{eff}} = \frac{1}{G r_{\mathrm{eff}}^2} \int_0^\infty (r - r_{\mathrm{eff}})^2 \pi r^2 n(r)\, \mathrm{d}r, \tag{5.142}$$

where G is the geometrical cross-section. We use the superscripts l and s to refer to the small and large mode of the size distribution, respectively.

B.2 Ocean description

For the retrieval simulations in this chapter, the lower boundary of the model atmosphere is characterized by the reflection matrix of the ocean. The ocean reflection can be described by three contributions (see, for example, Chowdhary (1999) and references therin): (1) Fresnel reflection on the oceanic waves.

This contribution is mainly determined by the wind speed W. (2) Scattering inside the ocean body called underlight. In this chapter we restrict ourselves to the open ocean (so called 'case-1 waters' (Morel and Prieur, 1977)) for which the reflection due to underwater scattering is predominantly determined by the concentration of phytoplankton and its derivative products, referred to as the oceanic pigment concentration C_{pig}. (3) Reflection by oceanic foam, which depends on the foam albedo A_{fm} (see, for example, Koepke (1984), Frouin et al. (1996) and Kokhanovsky (2004)) and the fraction of the ground pixel that is covered by foam, which depends on the wind speed. So, the total ocean reflection depends mainly on the wind speed, the oceanic pigment concentration, and the foam albedo.

For the simulations in this chapter, the Fresnel reflection on the waves is calculated using the method of Mishchenko and Travis (1997), assuming the wind speed dependent distribution of surface slopes proposed by Cox and Munk (1954). Here, we used a windspeed $W = 7$ m/s throughout this chapter. For the foam albedo A_{fm} we assume a fixed value of 0.2, which is close to the value proposed by Koepke (1984) for the visible spectral range. For the wind speed dependent fraction l_{fm} of the ground pixel that is covered by foam we use $l_{fm} = 2.95 \times 10^{-6}\, W^{3.52}$ (Monahan and O'Muircheartaigh, 1980).

The underlight contribution is described using a Lambertian albedo that depends on the oceanic pigment concentration, using the dependence given by Morel (1988) and Morel and Gentili (1993) (an improved model has been published by Morel and Maritorena (2001)), in combination with the data of Smith and Baker (1981). So, bi-directional effects and polarization are neglected using this simplified description of underlight. Since the underlight contribution is largest below 500 nm, the errors in the underlight contribution also will be largest for these wavelengths. However, the effect of errors in the ocean description on the intensity vector at the top of the atmosphere will be relatively small, since the atmospheric contribution to the intensity vector at the top of the atmosphere is much larger than the oceanic contribution at these wavelengths. The neglect of bi-directional effects can cause an error in the underlight contribution of roughly 20% directly above the ocean surface (Morel and Gentili, 1993; Chowdhary, 1999) but is in general smaller than 1.5% at the top of the atmosphere for wavelengths below 500 nm. The neglect of polarization in the underlight contribution causes for some geometries maximum errors of 1–2% in Stokes parameter Q at the top of the atmosphere for realistic ocean models (Chowdhary, 1999). We expect that the simplified description of the underlight contribution does not significantly affect the sensitivity study results of this chapter. However, for aerosol retrieval from real measurements it is worthwhile to consider a more advanced ocean description (Chowdhary et al., 2005). Also for the retrieval of aerosol properties over coastal waters a more advanced ocean description should be considered.

Acknowledgements

This research was supported by the Dutch User Support Programme 2001-2005 (USP) under project number EO-069 and an Ozone Monitoring SAF Visiting Scientist Grant.

References

Bell, G. I. and Glasstone, S., 1970. *Nuclear Reactor Theory*. Van Nostrand Reinhold Company, New York.

Box, M., Gerstl, S., and Simmer, C., 1988. Application of the adjoint formulation to the calculation of atmospheric radiative effects. *Beitr. Phys. Atmos.*, **61**, 303–311.

Carter, L. L., Horak, H. G., and Sandford, M. T., 1978. An adjoint Monte Carlo treatment of the equation of radiative transfer for polarized light. *J. Comp. Phys.*, **26**, 119–138.

Chandrasekhar, S., 1960. *Radiative transfer*. Dover Publications, Inc., New York.

Chowdhary, J., 1999. *Multiple scattering of polarized light in atmosphere-ocean systems: Application to sensitivity analyses of aerosol polarimetry*. Ph.D. thesis, Columbia University.

Chowdhary, J., Cairns, B., Mishchenko, M., Hobbs, P., Cota, G., Redemann, J., Rutledge, K., Holben, B., and Russel, E., 2005. Retrieval of aerosol scattering and absorption properties from photo-polarimetric observations over the ocean during the CLAMS experiment. *J. Atmos. Sci.*, **62**(4), 1093–1117.

Cox, C. and Munk, W., 1954. Statistics of the sea surface derived from sun glitter. *J. Mar. Res.*, **13**, 198–227.

de Haan, J., Bosma, P., and Hovenier, J., 1987. The adding method for multiple scattering calculations of polarized light. *Astron. Astrophys.*, **181**, 371–391.

Deirmendjian, D., 1969. *Electromagnetic Scattering on Spherical Polydispersions*. Elsevier, New York.

de Rooij, W. A. and van der Stap, C. C. A. H., 1984. Expansion of Mie scattering matrices in generalized spherical functions. *Astron. Astrophys.*, **131**, 237–248.

Domke, H., 1975. Fourier expansion of the phase matrix for Mie scattering. *Z. Meteorologie*, **25**, 357.

Frouin, R., Schwindling, M., and Deschamps, P., 1996. Spectral reflectance of sea foam in the visible and near-infrared: In situ measurements and remote sensing implications. *J. Geophys. Res.*, **101**, 14 361–14 371.

Gel'fand, I., Minlos, R., and Shapiro, Z., 1963. *Representations of the Rotation and Lorentz Groups and their Applications*. Pergamon Press, Oxford.

Hansen, J. E. and Travis, L. D., 1974. Light scattering in planetary atmospheres. *Space Sci. Rev.*, **16**, 527–610.

Hansen, P., 1992. Analysis of discrete ill posed problems by means of the L-curve. *SIAM Rev.*, **34**, 561–580.

Hansen, P. and O'Leary, D., 1993. The use of the L-curve in the regularization of discrete ill posed problems. *SIAM J. Sci. Comput.*, **14**, 1487–1503.

Hasekamp, O. and Landgraf, J., 2002. A linearized vector radiative transfer model for atmospheric trace gas retrieval. *J. Quant. Spectrosc. Radiat. Transfer*, **75**, 221–238.

Hasekamp, O., Landgraf, J., Hartmann, W., and Aben, I., 2004. Proposal for GOME-2 PMD wavelength band selection and the effect on aerosol retrieval. Techn. Rep. SRON-EOS/RP/04-002, SRON, Utrecht, The Netherlands.

Hasekamp, O. P. and Landgraf, J., 2005a. Linearization of vector radiative transfer with respect to aerosol properties and its use in satellite remote sensing. *Journal of Geophysical Research (Atmospheres)*, **110**, 4203.

Hasekamp, O. P. and Landgraf, J., 2005b. Retrieval of aerosol properties over the ocean from multispectral single-viewing-angle measurements of intensity and polarization: Retrieval approach, information content, and sensitivity study. *Journal of Geophysical Research (Atmospheres)*, **110**, 20 207.

Hovenier, J. W., 1969. Symmetry relationships for scattering of polarized light in a slab of randomly oriented particles. *J. Atmos. Sci.*, **26**, 488–499.

Hovenier, J. W. and van der Mee, C. V. M., 1983. Fundamental relationships relevant to the transfer of polarized light in a scattering atmosphere. *Astron. Astrophys.*, **128**, 1–16.

Koepke, P., 1984. Effective reflectance of oceanic whitecaps. *Appl. Opt.*, **23**, 1816–1824.

Koepke, P. and Hess, M., 1988. Scattering functions of tropospheric aerosols: The effect of nonspherical particles. *Appl. Opt.*, **27**, 2422–2430.

Kokhanovsky, A. A., 2004. Spectral reflectance of whitecaps. *Journal of Geophysical Research (Oceans)*, **109**, 5021.

Kuščer, I. and Ribarič, M., 1959. Matrix formalism in the theory of diffusion of light. *Opt. Acta*, **6**, 42–51.

Landgraf, J., Hasekamp, O., and Trautmann, T., 2002. Linearization of radiative transfer with respect to surface properties. *J. Quant. Spectrosc. Radiat. Transfer*, **72**, 327–339.

Landgraf, J., Hasekamp, O., Trautmann, T., and Box, M., 2001. A linearized radiative transfer model for ozone profile retrieval using the analytical forward-adjoint perturbation theory. *J. Geophys. Res.*, **106**, 27,291–27,306.

Landgraf, J., Hasekamp, O., van Deelen, R., and Aben, I., 2004. Rotational Raman scattering of polarized light in the Earth atmosphere: A vector radiative transfer model using the radiative transfer perturbation theory approach. *J. Quant. Spectrosc. Radiat. Transfer*, **87**, 399–433

Levenberg, K., 1944. A method for the solution of certain problems in least squares. *Quart. Appl. Math.*, **2**, 164–168.

Lewins, J., 1965. *Importance, the Adjoint Function*. Pergamon Press, Oxford, England.

Marchuk, G., 1964. Equation for the value of information from weather satellites and formulation of inverse problem s. *Cosmic Res.*, **2**, 394–409.

Marquardt, D., 1963. An algorithm for least squares estimation of nonlinear parameters. *SIAM J. Appl. Math.*, **11**, 431–441.

Mishchenko, M. and Travis, L., 1994. Light scattering by polydispersions of randomly oriented spheroids with sizes comparable to wavelengths of observation. *Appl. Opt.*, **33**, 7206–7225.

Mishchenko, M. I., Lacis, A. A., Carlson, B., and Travis, L., 1995. Nonsphericity of dust-like tropospheric aerosols: Implications for aerosol remote sensing and climate modeling. *Geophys. Res. Lett.*, **22**, 1077–1980.

Mishchenko, M. I. and Travis, L. D., 1997. Satellite retrieval of aerosol properties over the ocean using polarization as well as intensity of reflected sunlight. *J. Geophys. Res.*, **102**, 16,989–17,013.

Monahan, E. and O'Muircheartaigh, I., 1980. Optical power law description of oceanic whitecap coverage dependence on windspeed. *J. Phys. Oceanogr.*, **10**, 2094.

Morel, A., 1988. Optical modeling of the upper ocean in relation to its biogenous matter content (case I waters). *J. Geophys. Res.*, **93**, 10 749–10 768.

Morel, A. and Gentili, B., 1993. Diffuse reflectance of oceanic waters. II. Bidirectional effects. *Appl. Opt.*, **32**, 6864–6879.

Morel, A. and Maritorena, S., 2001. Bio-optical properties of oceanic waters: A reappraisal. *J. Geophys. Res.*, **106**, 7163–7180.

Morel, A. and Prieur, L., 1977. Ananlysis of variations in ocean color. *Limnol. Oceanogr.*, **19**, 591–600.

Morse, P. M. and Feshbach, H., 1953. *Methods of Theoretical Physics*. McGraw-Hill, New York.

Phillips, P., 1962. A technique for the numerical solution of certain integral equations of the first kind. *J. Assoc. Comput. Mach.*, **9**, 84–97.

Press, W., Teukolsky, S., Vetterling, W., and Flannery, B., 1992. *Numerical Recipes in FORTRAN, the Art of Scientific Computing*. Cambridge University Press, Cambridge.

Rodgers, C., 1976. Retrieval of atmospheric temperature and composition from remote measurements of thermal radiation. *Rev. Geophys.*, **14**, 609–624.

Rodgers, C., 2000. *Inverse Methods for Atmospheric Sounding: Theory and Practice*. World Sc., River Edge, NJ.

Rodgers, C. D. and Connor, B. J., 2003. Intercomparison of remote sounding instruments. *Journal of Geophysical Research (Atmospheres)*, **108**, 13.

Rozanov, V., Kurosu, T., and Burrows, J., 1998. Retrieval of atmospheric constituents in the UV-visible: A new quasi-analytical approach for the calculation of weighting functions. *J. Quant. Spectrosc. Radiat. Transfer*, **60**, 277–299.

Schulz, F. M., Stamnes, K., and Weng, F., 1999. VDISORT: An improved and generalized discrete ordinate method for polarized (vector) radiative transfer. *J. Quant. Spectrosc. Radiat. Transfer*, **61**, 105–122.

Sendra, C. and Box, M., 2000. Retrieval of the phase function and scattering optical thickness of aerosols: A radiative transfer perturbation theory application. *J. Quant. Spectrosc. Radiat. Transfer*, **64**, 499–515.

Smith, R. and Baker, K., 1981. Optical properties of the clearest natural waters. *Appl. Opt.*, **20**, 177–184.

Spurr, R., Kurosu, T., and Chance, K., 2001. A linearized discrete ordinate radiative transfer model for atmospheric remote-sensing retrieval. *J. Quant. Spectrosc. Radiat. Transfer*, **68**, 689–735.

Stammes, P., de Haan, J., and Hovenier, J., 1989. The polarized internal radiation field of a planetary atmosphere. *Astron. Astrophys.*, **225**, 239–259.

Stamnes, K., Tsay, S.-C., Wiscombe, W., and Jayaweera, K., 1988. Numerically stable algorithm for discrete-ordinate-method radiative transfer in multiple scattering and emitting layered media. *Appl. Optics*, **27**, 2502–2509.

Tikhonov, A., 1963. On the solution of incorrectly stated problems and a method of regularization. *Dokl. Akad. Nauk SSSR*, **151**, 501–504.

Torres, O., Bhartia, P. K., Herman, J. R., Ahmad, Z., and Gleason, J., 1998. Derivation of aerosol properties from satellite measurements of backscattered ultraviolet radiation: Theoretical basis. *J. Geophys. Res.*, **103**, 17 099–17 110.

Torres, O., Decae, R., Veefkind, P., and de Leeuw, G., 2001. OMI aerosol retrieval algorithm. *ATBD-OMI-03*, pages 47–69.

Ustinov, E. A., 1988. Methods of spherical harmonics: Application to the transfer of polarized radiation in a vertically non-uniform planetary atmosphere. Mathematical apparatus. *Cosmic Res.*, **26**, 473.

Ustinov, E. A., 1991. Inverse problem of photometric observation of solar radiation reflected by an optically dense planetary atmosphere. Mathematical methods and weighting functions of linearized inverse problem. *Cosmic Res.*, **29**, 519–532.

Ustinov, E. A., 1992. Inverse problem of the photometry of solar radiation reflected by an optically thick planetary atmosphere. 3. Remote sensing of minor gaseous constituents and an atmospheric aerosol. *Cosmic Res.*, **30**, 170–181.

Ustinov, E. A., 2001. Adjoint sensitivity analysis of radiative transfer equation: Temperature and gas mixing ratio weighting functions for remote sensing of scattering atmospheres in thermal IR. *J. Quant. Spectrosc. Radiat. Transfer*, **68**, 195–211.

van de Hulst, H. C., 1957. *Light Scattering by Small Particles*. John Wiley and Sons, New York.

Wiscombe, W. and Grams, G., 1988. Scattering from nonspherical Chebyshev particles, 2, Means of angular scattering patterns. *Appl. Opt.*, **27**, 2405–2421.