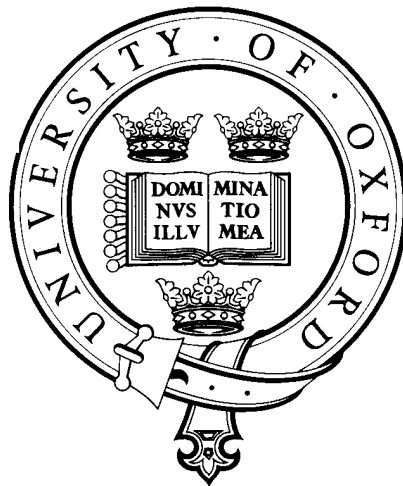


The Scattering of Light by Non-spherical Particles

Andrew John Alexander Smith

Supervisor: Don Grainger

Postdoctoral Advisers: Dan Peters, Elisa Carboni



First Year Report

Atmospheric, Oceanic and Planetary Physics

Department of Physics

University of Oxford

August 2008

Contents

1	Introduction	1
1.1	What are aerosols?	1
1.2	Sources	3
1.2.1	Natural aerosols	3
1.2.2	Anthropogenic aerosols	8
1.3	The influence of aerosols on the atmosphere	10
1.3.1	Radiative influence	10
1.3.2	Chemical influence	12
1.3.3	Clouds	12
1.4	Important aerosol optical characteristics	13
1.4.1	Size parameter	13
1.4.2	The amplitude matrix	13
1.4.3	Stokes parameters	14
1.4.4	The phase function	15
1.4.5	Asymmetry parameter	15
1.4.6	Cross-sections and efficiencies	15
1.4.7	Single scatter albedo	16
2	Non-spherical Scattering Methods	17
2.1	The spherical basis - Mie theory	17
2.1.1	The vector wave equations	18
2.1.2	The incident wave in spherical harmonics	19
2.1.3	The scattered field	20
2.1.4	Boundary conditions	21
2.1.5	The amplitude matrix, efficiency factors and Stokes parameters	22
2.2	T-matrix theory	23

2.2.1	Vector Spherical Wave functions	23
2.2.2	The amplitude matrix and other important physical properties	26
2.2.3	Orientation averaging	26
2.2.4	Calculating the T-matrix	27
2.3	Fractal representation	30
2.3.1	The scaling approach	31
2.4	Examples of Scattering calculations	32
3	SPARCLE Design and Optimisation	37
3.1	The forward model	37
3.1.1	Detector geometry	38
3.1.2	The phase function and its derivatives	40
3.2	Optimisation method	42
3.3	Conclusions	43
4	Time line and Future Work	48
4.1	Work to date	48
4.2	Plan for future work	49
4.2.1	Further work on SPARCLE	49
4.2.2	Research into additional scattering methods	49
4.2.3	Analytical differentials and the T-matrix method	50
4.2.4	Rapidity of calculations	51
	Bibliography	52
	A The expansion of vector spherical wave functions	57
	B The orthogonality of vector spherical wave functions	59

Abstract

The following is a report of my activities in the first year of my D.Phil in atmospheric physics. In the first chapter, there is a background to the subject of interest. The next chapter is a review of the relevant light scattering theory, including spherical and non-spherical scattering. In Chapter 3, work to optimise a particle detector is presented and finally a timetable and plan of future work is given.

Chapter 1

Introduction

1.1 What are aerosols?

An aerosol is a suspension of solid or liquid particles in a gas, generally with dimensions on length scales of nanometres to tens of microns. Initially, we can designate two categories: aerosols that arise from natural sources such as sand or emissions from volcanoes and those that are from anthropogenic sources. This second category would mainly consist of the many products of fuel combustion.

Primary aerosol is composed of particles emitted into the atmosphere. Secondary aerosol would be those particles formed in the atmosphere by gas-to-particle conversion processes [Seinfeld & Pandis 1998]. Concentrations can reach as high as several thousand per cubic centimetre for small particles ($< 1 \mu\text{m}$), whereas bigger particles have number densities less than 1 part. cm^{-3} .

Removing the particles from the atmosphere are two principal mechanisms: dry deposition (where the particles return to the Earth's surface) and wet deposition (the particles are incorporated into cloud droplets which can precipitate out).

Distribution

Differences between aerosol in the troposphere and stratosphere are very apparent. Tropospheric content can vary in size, shape (from spherical, to extremely non-spherical) and refractive index. By contrast, stratospheric aerosol is, for the most part, aqueous sulphuric acid of varying concentrations. Volcanic eruptions can be a large source of SO_2 which enters directly into the stratosphere, and, after these events, we see the stratospheric levels of sulphate gradually decrease back to a background level with an e-folding time of months [Mergenthaler et al. 1995].

Distribution of the particles is generally characterised by a log-normal distribution with three

principal modes. The smallest mode is nucleation (5×10^{-3} to $0.1 \mu\text{m}$), where particles are formed from the nucleation of atmospheric species in gas to particle conversion either on existing particles, or through condensation [Colbeck 1998].

The second, and most populated mode, is the accumulation mode with particle diameters of $.1$ to $2.5 \mu\text{m}$. Here, we find particles which arise from the condensation of low-volatility vapours. Particles in this range are subject to the least number of removal mechanisms, so numbers accumulate here. These first two modes are collectively named the fine mode.

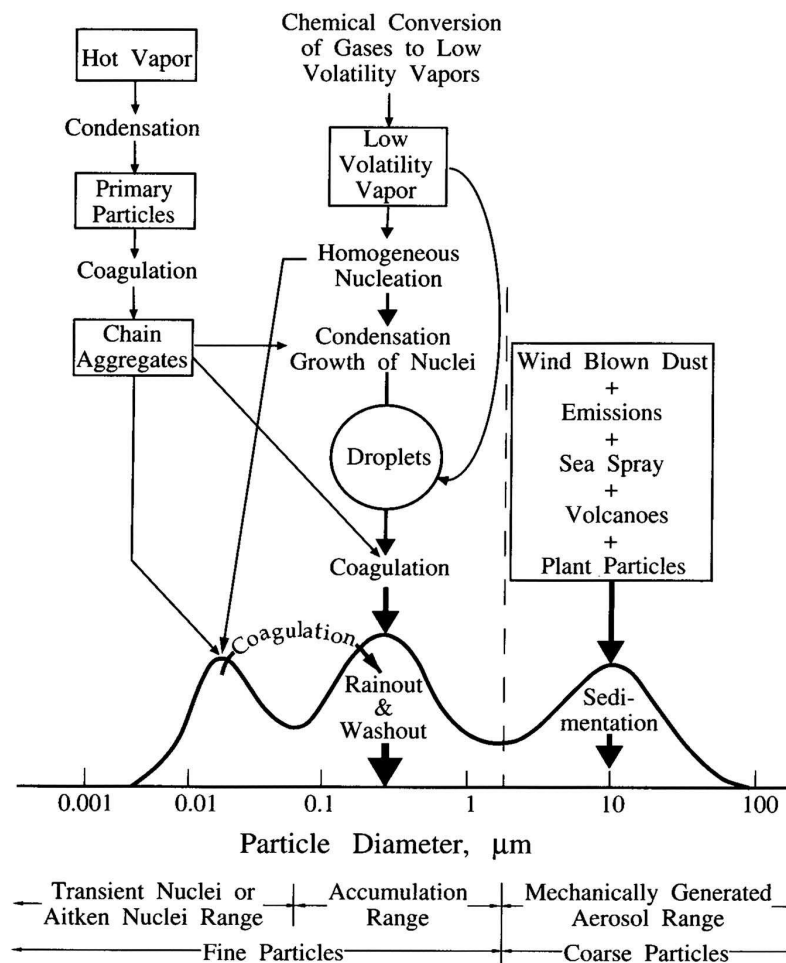


Figure 1.1: Idealised schematic of the distribution of particle surface area of atmospheric aerosol, including sources and sinks for the modes (from Chapter 7 of Colbeck [1998], reproduced from Whitby & Cantrell [1976]).

Finally, the coarse mode ($> 2.5 \mu\text{m}$) consists of large particles, formed by mechanical processes

such as weathering and wind erosion. Due to their size, they have high sedimentation velocities (which increase roughly quadratically with size) and settle out of the atmosphere reasonably quickly (by dry deposition). The sources and sinks for the fine and coarse modes are different, and so we find very little exchanging of aerosols between the two.

Fig. 1.1 shows the three principal modes and their methods of growth and decay. Notice the one-way increase in the accumulation mode from the nucleation mode and the lack of exchange from fine to coarse particles.

1.2 Sources

Unlike most atmospheric gases, many aerosols have mixing times which are shorter than their atmospheric lifetimes. This leads to spatially non-homogeneous distributions of individual species. Consider the huge smog problems of Los Angeles where we find high concentrations of particulates, tropospheric ozone and NO_x leading to low quality air and low visibility. Very steady meteorological patterns caused by the semi-permanent sub-tropical anticyclone over the Pacific ocean leads to descending air, high temperatures and very low wind speeds and inversion heights along the Californian coast. These conditions “trap” particulate in the cooler marine layer, where they are prevented from vertical exchange with the upper-air by the temperature inversion [Seinfeld & Pandis 1998].

We can also see the non-homogeneous distribution of aerosols in Fig. 1.2 which shows total aerosol optical depth for two periods in 2001. It is immediately obvious that certain areas of the globe have higher concentrations of aerosol and often these are the industrial, continental regions. Increased optical depth also marks a clear plume extending into the Atlantic ocean from Saharan Africa and biomass emissions over the Amazon rainforests.

The following sections give a brief outline of aerosol sources. Particular attention is paid to those which are aspherical. It would be expected that these particles would be inferred badly in retrieval schemes which used the assumption of sphericity. Carbon chains and sand cause the most problems because they are highly non-spherical. Although the categories “natural” and “anthropogenic” are not strictly accurate names¹, most species are created predominantly by one mechanism. The particulate that is *mainly* in the atmosphere due to human activity will be termed anthropogenic.

1.2.1 Natural aerosols

The sources of natural aerosols are controlled by local environmental conditions. Wind, temperature and humidity all play a part in effecting the rate of emission. Many of the particles are “kicked

¹For example, burning can also be because of natural wild fires.

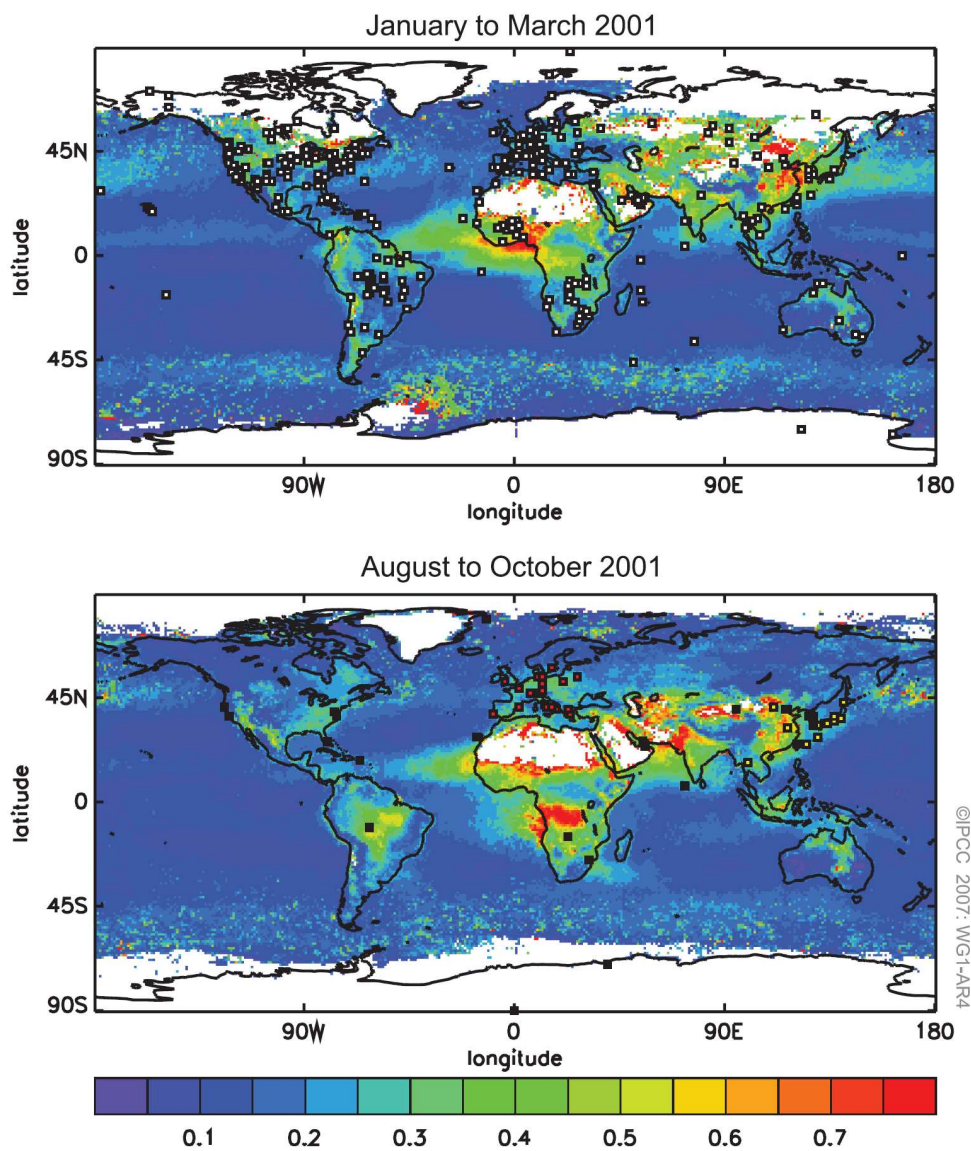


Figure 1.2: Total aerosol optical depth determined from satellite measurements, showing seasonal and geographical changes in aerosol concentration. In white areas, no values are available due to the high reflectivity of sand (in the Sahara) and snow (in the Arctic and Antarctic regions). This plot is Fig. TS.4. from the IPCC report [Solomon et al. 2007].

up” into the atmosphere so wind speed is the main factor determining population, and we see large temporal variation in natural aerosol. Of the natural aerosols, the significant non-spherical component is mineral dust.

Salt

Marine aerosol is mainly composed of sea-salt. At the surface of the ocean, air bubbles burst, spraying droplets which either fall back to the surface of the water (if they are too large) or are carried higher, forming aerosol. Depending on the wind speed, the aerosol will or will not be able to reach the boundary layer (the faster the wind, the more likely) and we find that sea-salt concentration varies roughly linearly with wind speed [Colbeck 1998]. Sea-salt aerosol is also responsible for sulphate formation (as an active site). The main salts are NaCl, KCl, CaSO₄ and Na₂SO₄ and atmospheric emission is estimated at 1300 Tg yr⁻¹ [Colbeck 1998, Table 7.2]. This is a large proportion of the natural tropospheric aerosol but we find concentrations rapidly decrease as we travel inland from the oceans. Salt aerosol can be considered spherical as it will generally be suspended in water droplets.

Mineral dust

Mineral dust from desert and semi-arid regions is a major source of tropospheric aerosol. About 50 % of this is subject to long-range transport and will turn up all over the world. Fig. 1.3 shows a dust plume extending out over the Atlantic ocean from the West coast of Africa; this is an example of the huge influence that mineral constituents can have on the optical properties of the atmosphere. Saharan sand is thought to play a vital role in the support of vegetation in the Amazon basin and provides nutrient iron to the biogeochemical cycle of the ocean systems [Solomon et al. 2007].

Dust is brought into the atmosphere by saltation, a jumping motion whereby already detached, larger particles collide with obstructions on the surface bed, projecting smaller particles into the air. Interestingly, wind on its own is not enough to remove particles from the surface bed. Experiments carried out by Bagnold in the 1940s found that when a steady stream of air was blown over loosely scattered, fine cement, no particle motion was recorded, even at very high wind speeds. Larger particles, already in motion, were required to lift the particulate. A good description of saltation is provided in Twomey [1977].

Since all of the fine sand (which is light enough to be lifted into the troposphere) has been removed from the surface, most of the Saharan region does not provide huge sources of dust and, generally, the sand is very coarse. Specific regions such as the Bodélé Depression in Chad, where annual deposition occurs, along with other preferable circumstances (including strong surface winds and favourable topography), provide the world with its sand, this in fact being the world’s largest dust source [Washington

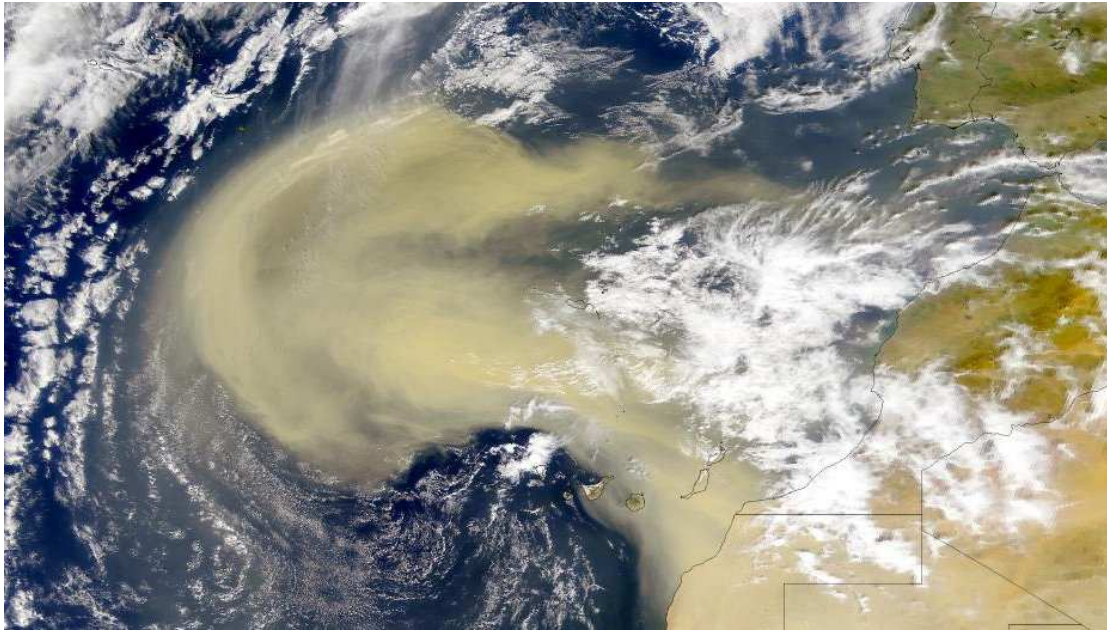


Figure 1.3: Sand storm event off the West African coast, 26th February, 2000. From *NASA Earth Observatory News Room* [2000]

et al. 2006]. Apart from the Sahara, the other major world source of mineral dust is the Gobi desert, with the US South West and the Saudi Arabian peninsula providing more local sources without the global reach.

The size of mineral dust in the atmosphere is around $1 \mu\text{m}$ and varies hugely in shape. Fig. 1.4 shows scanning electron microscopy of sample Saharan sand², demonstrating the sharp corners and varying shape which make these particles most certainly not spheres.

²It is highly likely that this sand sample was collected at ground level since the sizes are not representative of dust in the troposphere, being too large.

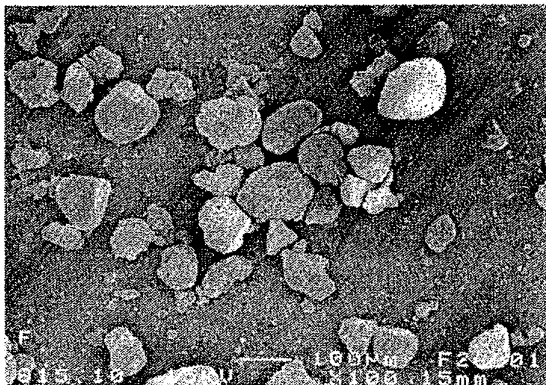


Figure 1.4: Scanning electron microscope photograph of Saharan sand. This is Fig. 2(g) from Volten et al. [2001].

Another class of mineral dust, soil enters the atmosphere in a similar way to sand, but comes from different regions and has different consistency. We also need to consider the moisture of the surface, and the vegetation cover, both of which can prevent removal of particulate.

Principal constituents of mineral dust are oxides and carbonates of silicon, aluminium, calcium and iron. Table 1.1 gives the abundances of major soil constituent elements.

Si	Al	Fe	Ca	Mg	Na	K	Ti	Mn	Cr	V	Co
330,000	71,300	38,000	13,700	6,300	6,300	13,600	4,600	850	200	100	8

Table 1.1: Average abundances of major elements in soil, from Seinfeld & Pandis [1998]. Elemental abundances are in ppm by mass.

Volcanic emissions

Solid and gaseous material is injected into the atmosphere sporadically with the eruption of volcanoes. The larger solids are heavy and mostly deposited in the local area due to gravity, but huge amounts of ash, SO_2 and HCl are left in the atmosphere. Ash itself principally consists of SO_2 , Al_2O_3 and Fe_2O_3 and is in the coarse aerosol mode, mostly impacting on the regional scale. Globally, the main result of volcanoes is the creation of sulphate aerosol in the stratosphere. The explosive nature of eruptions means that sulphate can be injected buoyantly into the stratosphere where there are none of the precipitation effects that make sulphate lifetimes in the troposphere so much shorter. Thus residence times for stratospheric aerosols is estimated at 6–9 months [Colbeck 1998].

The eruption of Mount Pinatubo on Luzon Island, Philippines in 1991 was the largest in recent times. Satellite measurements of column SO_2 estimated 20,000 kt was released and encircled the Earth within 22 days [Bluth et al. 1992]. As an example of the impact of these events, consider the year of the eruption. Records show that the warmest years in the 20th century have all taken place after the late 1960s. In 1991 and 1992 there was a sudden drop in global mean air temperature, due to the spread of aerosols in the stratosphere from the Pinatubo eruption [Bhutiyan et al. 2007], the only significant blip in an otherwise upward trend of global temperature. It took more than two years for aerosol quantities in the stratosphere to decay back to “normal” levels.

Since the sulphate aerosol is dissolved H_2SO_4 , it can be assumed to be spherical. This is not the case for the ash, but we generally find that in periods not directly after an eruption there is not much of this in the air.

1.2.2 Anthropogenic aerosols

Anthropogenic aerosols come from particulate emissions from industry, transport and natural sources such as farming. In urban areas, the spread of sizes can be very variable, with sample sites close to a source (for example, a road) having extremely high concentrations of fine particulate. As we move away from the source, we see a rapid decrease in the concentrations of these smaller particles, as shown in Fig. 1.5.

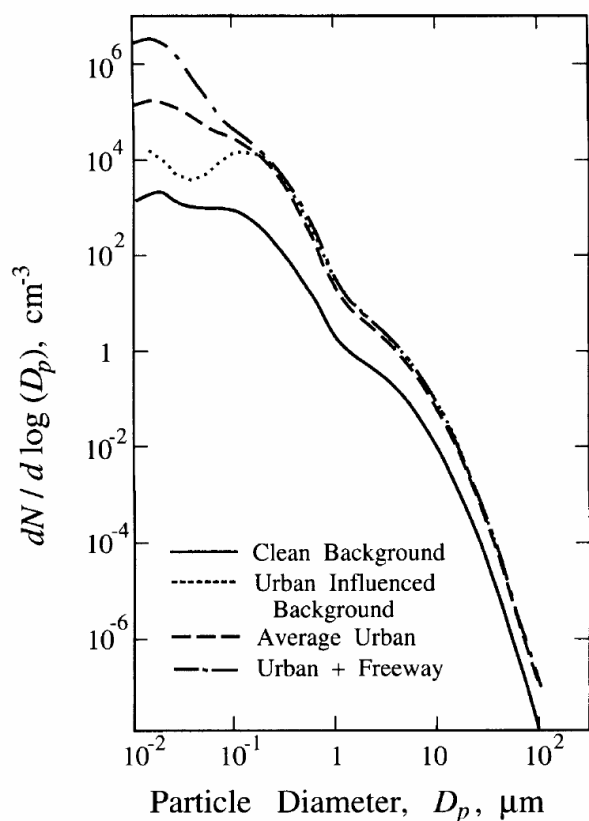


Figure 1.5: Comparing aerosol number next to sources, and for average urban background. Reproduced from Seinfeld & Pandis [1998].

The combustion of fuels also leads to secondary material from the chemical reactions in burning. These are gas to particle reactions in the atmosphere and the products are sulphates, nitrates and ammonium. Industry and agriculture are the main sources with mining, stone crushing in quarries and grain elevators all contributing heavily to coarse mode particles. However, the main area of interest from a non-spherical scattering perspective is black carbon, and other, secondary organics which are highly non-spherical.

Carbon

The main carbon molecules referred to are black carbon (BC), but organic molecules and a plethora of smaller carbon molecules created by combustion should also be mentioned.

The light-scattering properties of BC aggregates are tricky characters to investigate. Unfortunately, BC is one of the most prevalent anthropogenic aerosol. Table 7.1 from Chapter 7 of Colbeck [1998] shows that of all compounds and elements in the urban aerosol class, 29% by mass are organics ($30.4 \mu\text{g m}^{-3}$ of total $103.6 \mu\text{g m}^{-3}$). In fact, during the burning season, one third of the South African continent can be covered in smoke aerosol from biomass burning [Remer et al. 1998]. It should be mentioned that not all of the carbon in the atmosphere is from anthropogenic sources. Natural emissions from the biosphere also contribute to this group.

Biomass burning is the first significant source of emissions. Again, some of these are from natural wild fires, but the largest contributor is the clearing of forest and savannah for agricultural purposes. The particles are incomplete combustion products, soot, sulphate, nitrate and hydrocarbons. Gases, CO_2 , CO , CH_4 and volatile organic compounds are also emitted. The type of fuel burnt makes a difference. Spark-ignition engines have high organic carbon content, while diesel engines emit more elemental carbon [Colbeck 1998].

BC chains undergo a rapid evolution in the atmosphere due to coagulation, outgassing, condensation and gas-particle conversion. Starting as long, open chains, soot will gradually collapse into densely packed clusters over a time period of hours [Abel et al. 2003]. Reid & Hobbs [1998] found that when measuring biomass aerosol very close to source, the single scatter albedo, $\bar{\omega}_0$, could be very low compared to the evolved chains which had spent more time in the atmosphere. Combustion with high BC content (30%) could have $\bar{\omega}_0$ at the limit set by Mie theory for highly absorbing particles with similar size scales. Knowledge of refractive index and particularly shape are not well known. Other properties such as mixing type of the suspension can result in differing absorbing of light so a better picture of these particles is very important. We can also find that the particles become coated by a non-absorbing shell such as organics or sulphates (forming an internal mixture) if the BC particles are formed at a high temperature [Martins et al. 1998].

A useful property of soot is that it is very persistent in the atmosphere, due to being inert and relatively hydrophobic. This makes it a good tracer for anthropogenic aerosol emission.

Secondary anthropogenic emissions

The main source of secondary emissions is oxidation of SO_2 and the various nitrous oxides (NO_x). About 50% of these species will undergo oxidation before being deposited. Atmospheric OH^- radicals oxidise to H_2SO_4 and HNO_3 . Sulphuric acid has low vapour pressure, and all of this thus becomes

aerosol mass. Nitric acid is split between gas and aerosol phases. The SO_2 comes from fossil fuel emissions and is the major component of upper tropospheric and stratospheric aerosol. The secondary aerosol are in the fine mode and spherical.

1.3 The influence of aerosols on the atmosphere

Aerosol influence on the atmosphere is, at the time of writing, the largest unknown in estimates of radiative forcing. Fig. 1.6 shows the IPCC's³ best estimate for global mean radiative forcings in 2005, with their 90% confidence intervals marked. We can see immediately that by far the widest confidence intervals are for the aerosol effects and that there is a low level of scientific understanding. This gives an important motivation for research into aerosol light scattering properties.

Aerosols reflect light back into space and act as cloud condensation nuclei (CCN), increasing the albedo. They also act as surface for other chemical reactions in the atmosphere to take place, for example the destruction of ozone in the troposphere and the nighttime formation of nitric acid. Their temporal variability, different types, and the number of different effects and reactions they influence makes them fiendishly difficult to model.

1.3.1 Radiative influence

Radiative forcing (RF) is a measure of how changes to the atmosphere affect the climate system. It can be defined as the net change in irradiance at the tropopause. A positive value implies that there is more incoming than outgoing radiation⁴. In Fig. 1.6, we see that the aerosol RF has been split into two categories, the direct effect and the cloud albedo effect.

The direct effect is the scattering and absorption of light, altering the Earth's radiative balance. If the aerosols are purely scattering, the effect is always a negative RF since some of the light destined to pass through an atmospheric layer is now reflected, and will pass back into space. For absorbing aerosols, the effect is more variable. Haywood & Shine [1995], using a simplified formula, show that radiative forcing will be negative only if the single scatter albedo, $\bar{\omega}_0$ satisfies:

$$\bar{\omega}_0 > \frac{2R_s}{\bar{\beta}(1 - R_s)^2 + 2R_s}, \quad (1.1)$$

where R_s is the surface reflectance and $\bar{\beta}$ is the backscatter fraction, the ratio of backscattered to

³Intergovernmental Panel on Climate Change. See Solomon et al. [2007]

⁴The IPCC defines RF slightly differently as the “rate of energy change per unit area of the globe as measured at the top of the atmosphere” since 1750 (the beginning of the industrial era).

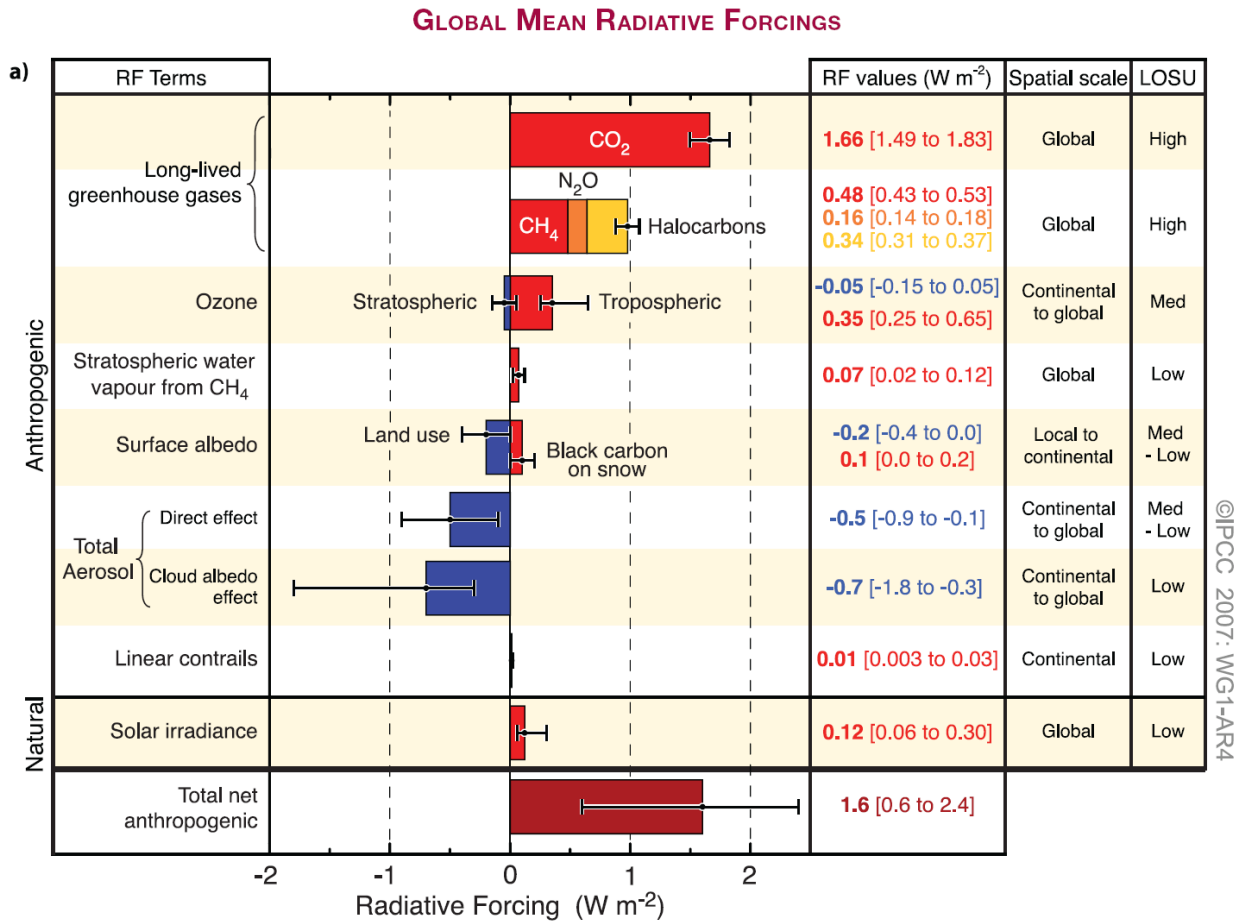


Figure 1.6: The estimates for global mean radiative forcings, presented by Working Group One of the IPCC. LOSU is the level of scientific understanding. Notice that the two largest error bars are those on the estimates of the aerosol direct and indirect effects. This is Fig. TS.5 in Solomon et al. [2007].

forward scattered light, obtained from the phase function [Wiscombe & Grams 1976]. Hence we find that over dark surfaces (forests or oceans), the effect is still negative while over bright surfaces (desert or snow or if the aerosol is above a cloud), there can be positive RF.

Short wavelength light rays are affected by all aerosol, leading to less irradiance at ground level. Long wavelength forcing is only appreciable if the scattering particles are large enough, and so has less of an effect.

The indirect aerosol effect refers to interaction with clouds and will be discussed in §1.3.3.

1.3.2 Chemical influence

Heterogeneous reactions involve the reaction of a gas molecule with a solid or liquid surface. This makes aerosol ideal atmospheric reactants in these cases.

In the troposphere for example, dry sea-salt acts to liberate hydrogen chloride when it comes into contact with gaseous nitric acid. Ozone destruction can also take place on the surface of carbon based aerosol, leaving us with CO and CO₂. Hazes in the troposphere offer wet surfaces for heterogeneous reactions, for example the nighttime formation of nitric acid.

In the stratosphere, polar stratospheric clouds, consisting of frozen or liquid sulphuric and nitric aerosol, are responsible for heterogeneous chlorine activation reactions which have a heavy role in polar ozone depletion [Colbeck 1998].

1.3.3 Clouds

The indirect aerosol effect consists of several interactions which change the microphysical properties of clouds. Aerosols are vital in cloud formation, since they act as cloud condensation nuclei (CCN), and since clouds are a big influence on the Earth's albedo due to their high reflectivity, changes in atmospheric aerosol concentration inevitably have an effect on RF (although, as Fig. 1.6 has demonstrated, an understanding of the subtleties is in no way complete).

The first indirect effect, or Twomey effect [Twomey 1977], is an obviously direct consequence of increased particulate in the atmosphere: more CCN mean that there are more sites where droplets can form in developing clouds. This leads to greater numbers of droplets, but of a smaller size than in an unperturbed cloud (leading to an increase in the surface area of water droplets in the cloud). The end result for the radiative properties of the cloud is an increase in light reflected back into space and thus a negative radiative forcing.

Second indirect effects include the suppression of drizzle with greater numbers of CCN and other factors [Albrecht 1989, Yia et al. 2008] which increase the liquid water content of the cloud. The cloud height can also increase, which leads to radiation as the clouds are warmed up [Pincus & Baker 1994]. The lifetime of the cloud is also increased since the reduced drizzle stops the cloud from breaking up [Albrecht 1989].

A good overview of cloud effects can be found in the IPCC report [Solomon et al. 2007].

1.4 Important aerosol optical characteristics

1.4.1 Size parameter

When looking at how light will scatter from a particle, the most important question is how comparable is the wavelength of light to the particle's size. If the wavelength is much greater than the size of the scattering object, the object won't be able to deflect the light in any significant way. As a result, the size parameter, x , is a clearer measure of the size regime in which we are interested than an absolute measure of the particle.

$$x = \frac{2\pi}{\lambda} r_0 = k r_0, \quad (1.2)$$

where λ is the wavelength of light and r_0 is the characteristic length of a particle and k is the circular wavenumber. For a sphere, r_0 , is the radius of the particle but for non-spherical particles, the choice of r_0 is not so clear. Mishchenko & Travis [1998] have written T-matrix code which gives users the option of defining r_0 as the radius of a sphere of either equal volume, or equal surface area to that of the scattering object.

1.4.2 The amplitude matrix

The amplitude matrix relates the scattered amplitude, \mathbf{E}_s , to the incident amplitude of the electric field, \mathbf{E}_i , where the light is split into light polarised parallel and perpendicular to the plane of scattering.

$$\begin{aligned} \mathbf{E}_s &= \frac{e^{ikr}}{k r} \mathbf{S} \mathbf{E}_i, \\ \begin{pmatrix} E_{\parallel s} \\ E_{\perp s} \end{pmatrix} &= \frac{e^{ikr}}{k r} \begin{pmatrix} S_2 & S_3 \\ S_4 & S_1 \end{pmatrix} \begin{pmatrix} E_{\parallel i} \\ E_{\perp i} \end{pmatrix}. \end{aligned} \quad (1.3)$$

Definitions vary slightly between texts. For example, Bohren & Huffman [1983] multiply the prefactor by i , or Mishchenko et al. [2000] by k (as well as using $E_\theta = E_{\parallel}$ and $E_\phi = -E_{\perp}$). Note the somewhat odd choice of element numbering in (1.3) which is due to convention in solving for the spherical case (see Mie theory, §2.1). This is a far field solution where it can be assumed that the scattered field is transverse ($\mathbf{E}_s \cdot \hat{\mathbf{e}}_r \simeq 0$). The elements of \mathbf{S} will be complex.

1.4.3 Stokes parameters

Stokes parameters are closely linked to the electric field, but have been chosen to mirror easily measurable scattering quantities. From van de Hulst [1957], we say:

$$\mathbf{I} = \begin{pmatrix} I \\ Q \\ U \\ V \end{pmatrix} \quad \text{where} \quad \begin{aligned} I &= E_{\parallel} E_{\parallel}^* + E_{\perp} E_{\perp}^*, & Q &= E_{\parallel} E_{\parallel}^* - E_{\perp} E_{\perp}^*, \\ U &= E_{\parallel} E_{\perp}^* + E_{\perp}^* E_{\parallel}, & V &= i(E_{\parallel} E_{\perp}^* - E_{\perp}^* E_{\parallel}). \end{aligned} \quad (1.4)$$

The intensity is I and the other elements measure the various amounts of polarisation. For a completely polarised beam, we have $I^2 = Q^2 + U^2 + V^2$ and for partially polarised beams, $I^2 \geq Q^2 + U^2 + V^2$. The degree of linear polarisation is given by Q and U and circular polarisation by V , as shown in Table 1.2.

Direction of polarisation for unit intensity ($I = 1$)	Stokes parameter		
	Q	U	V
Unpolarised	0	0	0
$\hat{\mathbf{e}}_{\parallel}$	1	0	0
$\hat{\mathbf{e}}_{\perp}$	-1	0	0
$(\hat{\mathbf{e}}_{\parallel} + \hat{\mathbf{e}}_{\perp})/\sqrt{2}$	0	1	0
$(\hat{\mathbf{e}}_{\parallel} - \hat{\mathbf{e}}_{\perp})/\sqrt{2}$	0	-1	0
Right circular	0	0	1
Left circular	0	0	-1

Table 1.2: Showing the meaning of the values of different Stokes Parameters for polarised light. Adapted from Bohren & Huffman [1983].

Scattered parameters can be calculated from the incident parameters using the phase matrix, \mathbf{Z} :

$$\begin{pmatrix} I_s \\ Q_s \\ U_s \\ V_s \end{pmatrix} = \frac{1}{k^2 r^2} \begin{pmatrix} Z_{11} & Z_{12} & Z_{13} & Z_{14} \\ Z_{21} & Z_{22} & Z_{23} & Z_{24} \\ Z_{31} & Z_{32} & Z_{33} & Z_{34} \\ Z_{41} & Z_{42} & Z_{43} & Z_{44} \end{pmatrix} \begin{pmatrix} I_i \\ Q_i \\ U_i \\ V_i \end{pmatrix}. \quad (1.5)$$

The values for elements Z_{ij} (as a function of the amplitude matrix) are (3.16) in Bohren & Huffman [1983] and (13–29) in Chapter 1 of Mishchenko et al. [2000]. As with the definition of the amplitude matrix, slight differences occur. Mishchenko et al. [2000] multiply (1.5) by k^2 and have differing signs

for some off diagonal elements, due to the use of unit vector $\hat{\mathbf{e}}_\phi$ instead of $\hat{\mathbf{e}}_\perp$.

1.4.4 The phase function

The phase function is a normalised measure of how the intensity of scattered light varies with scattering angle. For a single particle, with unpolarised light, this is defined as [van de Hulst 1957]:

$$p(\theta) = \frac{2}{C_{\text{sca}} k^2 r^2} (|S_1(\theta)|^2 + |S_2(\theta)|^2), \quad (1.6)$$

where C_{sca} is the scattering cross-section which will be defined in §1.4.6, and the intensity can be calculated as $I = I_0 C_{\text{sca}} p(\theta)$.

1.4.5 Asymmetry parameter

The asymmetry parameter is calculated from the phase function and tells us whether more light scatters in the forward or backwards direction.

$$g = \langle \cos \theta \rangle = \int_{4\pi} p(\theta) \cos \theta d\Omega. \quad (1.7)$$

Positive values indicate more forwards scattering. If $g = 0$ then light scatters equally in all directions (or symmetrically about a scattering angle of $\theta = 90$).

1.4.6 Cross-sections and efficiencies

Cross-sections tell us how much of an incident beam's intensity is taken away⁵ in terms of the area of the beam. If the power taken away by a certain process is W_x , then the cross-section for this process, C_x , is:

$$C_x = \frac{W_x}{\frac{1}{2} \sqrt{\frac{\epsilon}{\mu}} |\mathbf{E}_{i0}|^2}.$$

For example, scattering cross-section, C_{sca} , tells us what area of the incident beam becomes scattered light. Extinction cross-section, C_{ext} , tells us how much of the incident light is removed from the incident beam by the particle. The absorption cross-section, C_{abs} , tells us how much of the incident light is absorbed.

⁵The cross section is taken in a plane perpendicular to the direction of incident beam.

$$C_{\text{sca}} = \frac{1}{I_i} \int_{4\pi} I_s(\mathbf{n}_s) d\mathbf{n}_s, \quad (1.8)$$

$$C_{\text{ext}} = \frac{2\pi}{k} \Im \{Z_{11} + Z_{22}\} \quad (1.9)$$

$$C_{\text{abs}} = C_{\text{ext}} - C_{\text{sca}} \quad (1.10)$$

where $I_s = Z_{11} I_i + Z_{12} Q_i + Z_{13} U_i + Z_{14} V_i$ and Z_{ij} can be obtained from the Stokes parameters defined in (1.5).

Efficiencies are closely linked to cross-sections, but tell us what *proportion* of the beam incident on a particle is diverted to a certain process. As such, we define efficiencies, Q , as:

$$Q_{\text{ext}} = \frac{C_{\text{ext}}}{G}, \quad Q_{\text{sca}} = \frac{C_{\text{sca}}}{G}, \quad Q_{\text{abs}} = \frac{C_{\text{abs}}}{G}, \quad (1.11)$$

where G is the cross sectional area of a particle, projected onto a plane perpendicular to the direction of incident light. In the case of a sphere of radius r_0 , $G = \pi r_0^2$. Naively, we could guess that the extinction cross-section would be equal to G , but since the influence of the particle outside the limits of its physical shape is non-zero, this is not the case.

1.4.7 Single scatter albedo

The single scatter albedo tells us the probability that a photon incident on a volume element will survive.

$$\bar{\omega}_0 = \frac{C_{\text{sca}}}{C_{\text{ext}}}. \quad (1.12)$$

The range of values is $0 \leq \bar{\omega}_0 \leq 1$.

Chapter 2

Non-spherical Scattering Methods

Since the more rigorous methods below will rely on spherical harmonics, I will first outline the principal points of Mie theory, first published a century ago [Mie 1908], in §2.1. The T-matrix method is a popular, fast method of calculating light scattering properties, first by expressing the scattered fields as a matrix multiplied by the expansion of spherical harmonics. A brief explanation will be given in §2.2. Fractal representation defines an aggregate by a single, fractal dimension that relates the number of particles in an aggregate to the radius of the particle. The most basic fractal representation will be described in §2.3.

Over the course of the year, I have written IDL DLMS (dynamically loadable modules) to call Michael Mishchenko's FORTRAN implementation of T-matrix code for spherically symmetrical particles in IDL. The final section, §2.4, will show plots from this code, comparing spherical and non-spherical scatterers.

2.1 The spherical basis - Mie theory

Mie theory was developed by Gustav Mie in 1908 in order to understand the colours that resulted from light scattering from gold particles suspended in water. Although it is exact, only with the emergence of computing has it become practical to calculate the scattering values. An excellent introduction to the theory is provided in Bohren & Huffman [1983], whose notation and outline I have used. Code to calculate Mie scattering can be found on the AOPP website [*Light scattering routines* 2008].

2.1.1 The vector wave equations

For a linear, isotropic, homogeneous, divergence-free medium, we can obtain (from Maxwell's equations) the vector wave equation:

$$\begin{aligned}\nabla^2 \mathbf{E} + k^2 \mathbf{E} &= 0, & \nabla^2 \mathbf{H} + k^2 \mathbf{H} &= 0, \\ \nabla \cdot \mathbf{E} &= 0, & \nabla \cdot \mathbf{H} &= 0,\end{aligned}$$

where $k^2 = \omega^2 \epsilon \mu$, and,

$$\nabla \times \mathbf{E} = i\omega \mu \mathbf{H}, \quad \nabla \times \mathbf{H} = -i\omega \epsilon \mathbf{E}. \quad (2.1)$$

One can construct a vector function $\mathbf{M} = \nabla \times (\mathbf{c}\psi)$, where \mathbf{c} is constant and ψ is any scalar function. Since $\nabla \cdot (\nabla \times \mathbf{v}) = 0$ for any vector function \mathbf{v} , we can say that

$$\nabla \cdot \mathbf{M} = 0. \quad (2.2)$$

With rearrangement in the vector wave equation,

$$\nabla^2 \mathbf{M} + k^2 \mathbf{M} = \nabla \times [\mathbf{c} (\nabla^2 \psi + k^2 \psi)], \quad (2.3)$$

and thus, if ψ is a solution to the scalar wave equation, then \mathbf{M} is a solution to the vector wave equation. We can define a second vector function which obeys all of the above equations for \mathbf{M} ,

$$\mathbf{N} = \frac{\nabla \times \mathbf{M}}{k} \quad (2.4)$$

Now, changing our arbitrary constant vector \mathbf{c} to the radius vector, \mathbf{r} , we have solutions to the vector wave equation in spherical polar coordinates. Final forms of \mathbf{M} and \mathbf{N} are given in Appendix A.

Seeking separable solutions, we say $\psi(r, \theta, \phi) = R(r) \Theta(\theta) \Phi(\phi)$, and obtain the well known separated equations:

$$\frac{d^2 \Phi}{d\phi^2} = 0, \quad (2.5a)$$

$$\frac{1}{\sin \theta} \frac{d}{d\theta} \left(\sin \theta \frac{d\Theta}{d\theta} \right) + \left[n(n+1) - \frac{m^2}{\sin^2 \theta} \right] \Theta = 0, \quad (2.5b)$$

$$\frac{d}{dr} \left(r^2 \frac{dR}{dr} \right) + [k^2 r^2 - n(n+1)] R = 0, \quad (2.5c)$$

where $m = 0, 1, \dots$ and $n = 0, \dots, m$.

Following Bohren & Huffman [1983], we end up with even and odd scalar solutions:

$$\psi_{emn} = \cos m\phi P_n^m(\cos\theta)z_n(kr), \quad (2.6a)$$

$$\psi_{omn} = \sin m\phi P_n^m(\cos\theta)z_n(kr), \quad (2.6b)$$

and final vector solutions, \mathbf{M}_{emn} , \mathbf{M}_{omn} , \mathbf{N}_{emn} and \mathbf{N}_{omn} obtained from (2.3) with $\mathbf{c} = \mathbf{r}$. The radial function, $z_n(kr)$, is any linear combination of the spherical Bessel functions of first and second kind, $j_n(\rho)$ and $y_n(\rho)$. We will require the outgoing spherical Hankel function,

$$h_n^{(1)}(\rho) = j_n(\rho) + iy_n(\rho). \quad (2.7)$$

The zenith functions, $P_n^m(\cos\theta)$, are associated Legendre functions of which more can be found (along with information on Bessel functions) in Boas [1983].

2.1.2 The incident wave in spherical harmonics

The first stage is to correctly characterise the incident light. This is not as simple as it sounds, since it requires the expressing of a plane wave in vector spherical wave functions (VSWF).

An x-polarised wave can be written as

$$\begin{aligned} \mathbf{E}_i &= E_0 e^{ikr \cos\theta} \hat{\mathbf{e}}_x, \\ &= E_0 e^{ikr \cos\theta} (\sin\theta \cos\phi \hat{\mathbf{e}}_r + \cos\theta \cos\phi \hat{\mathbf{e}}_\theta + \sin\phi \hat{\mathbf{e}}_\phi). \end{aligned} \quad (2.8)$$

and expanded in VSWFs:

$$\mathbf{E}_i = \sum_{m=0}^{\infty} \sum_{n=m}^{\infty} (B_{emn} \mathbf{M}_{emn} + B_{omn} \mathbf{M}_{omn} + A_{emn} \mathbf{N}_{emn} + A_{omn} \mathbf{N}_{omn}). \quad (2.9)$$

where we need to find A_{emn} , A_{omn} , B_{emn} and B_{omn} . All of the VSWFs are mutually orthogonal, and it can be shown [Bohren & Huffman 1983] that $A_{omn} = B_{emn} = 0$ for all m, n . Coefficients are found using this orthogonality by:

$$A_{emn} = \frac{\int_0^{2\pi} \int_0^\pi \mathbf{E}_i \cdot \mathbf{N}_{emn} \sin\theta \, d\theta \, d\phi}{\int_0^{2\pi} \int_0^\pi |\mathbf{N}_{emn}|^2 \sin\theta \, d\theta \, d\phi}, \quad (2.10)$$

and similarly for B_{omn} . Due to the $\sin\phi$ and $\cos\phi$ found in (2.8), only the terms with $m = 1$ survive. If there was no scattering object at the origin, the wave would have to be finite, so we reject Bessel

functions of the 2nd kind, $y_n(\rho)$, since they have singularities here, using only $j_n(\rho)$. We write the VSWFs as $\mathbf{M}_{o1n}^{(1)}$ and $\mathbf{N}_{e1n}^{(1)}$, the ⁽¹⁾ indicating that the radial components are $z_n(kr) = j_n(kr)$.

After several hours of maths, we end up with the expansion coefficients:

$$A_{e1n} = E_0 i^{n-1} \frac{2n+1}{n(n+1)} \quad (2.11a)$$

$$B_{o1n} = E_0 i^n \frac{2n+1}{n(n+1)}, \quad (2.11b)$$

and an incident wave

$$\mathbf{E}_i = E_0 \sum_{n=1}^{\infty} i^n \frac{2n+1}{n(n+1)} \left(\mathbf{M}_{o1n}^{(1)} - i \mathbf{N}_{e1n}^{(1)} \right). \quad (2.12)$$

The magnetic field can then be simply obtained from (2.1).

2.1.3 The scattered field

We place a spherical medium at the origin, with radius, r_0 and refractive index m . In the same way that we expanded the incident field, \mathbf{E}_i , we can expand the scattered and internal fields, \mathbf{E}_s and \mathbf{E}_I . The form of the incident light and the orthogonality of the VSWFs determine the form of these. As with the incident field, the internal field can have no singularities and must also be composed only of Bessel functions of the first kind. The internal field also has a different value of wave vector, since the refractive index has changed. If we use the size parameter, $x = kr_0$, we can say that in the scattering medium, the wave vector can be written as $k_I = k(N_I/N_0) = km$ so that the effective size parameter becomes mx . When working with the magnetic field, the value of μ will also change.

The scattered field does not have to worry about the origin so is allowed discontinuities there. It can be shown that the correct radial function is $h_n^{(1)}$ [Bohren & Huffman 1983]. The VSWFs for $R_n(r) = h_n^{(1)}(kx)$ are $\mathbf{M}_{o1n}^{(3)}$ and $\mathbf{N}_{e1n}^{(3)}$.

We thus find ourselves with four more unknowns, a_n , b_n , c_n and d_n for the two VSWFs in \mathbf{E}_s and \mathbf{E}_I . Writing

$$E_n = E_0 i^n \frac{2n+1}{n(n+1)}$$

we have

$$\mathbf{E}_s = \sum_{n=1}^{\infty} E_n \left(i a_n \mathbf{N}_{e1n}^{(3)} - b_n \mathbf{M}_{o1n}^{(3)} \right), \quad (2.13)$$

$$\mathbf{E}_I = \sum_{n=1}^{\infty} E_n \left(c_n \mathbf{M}_{o1n}^{(1)} - i d_n \mathbf{N}_{e1n}^{(1)} \right). \quad (2.14)$$

2.1.4 Boundary conditions

The boundary conditions imposed are:

$$(\mathbf{E}_i + \mathbf{E}_s - \mathbf{E}_I) \times \hat{\mathbf{e}}_r = (\mathbf{H}_i + \mathbf{H}_s - \mathbf{H}_I) \times \hat{\mathbf{e}}_r = 0 \quad (\text{for } r = r_0). \quad (2.15)$$

For the E-field, this leads to

$$\sum_{n=1}^{\infty} E_n \left[\mathbf{M}_{o1n}^{(1)} - b_n \mathbf{M}_{o1n}^{(3)} - c_n \mathbf{M}_{o1n}^{(1)} - i \left(\mathbf{N}_{e1n}^{(1)} - a_n \mathbf{N}_{e1n}^{(3)} - d_n \mathbf{N}_{e1n}^{(1)} \right) \right] = 0 \quad (\text{for } r = r_0). \quad (2.16)$$

Using the expansions of $\mathbf{M}_{o1n}^{(1)}$ etc. explicitly stated in (A.2), and defining functions

$$\pi_n(\theta) = \frac{P_n^1(\cos \theta)}{\sin \theta} \quad \text{and} \quad \tau_n(\theta) = \frac{dP_n^1(\cos \theta)}{d\theta},$$

we can look at the angular components of (2.15) separately to find (for \mathbf{E}):

$$\begin{aligned} \hat{\mathbf{e}}_\theta : \quad & \pi_n [j_n(x) - b_n h_n^{(1)}(x) - c_n j_n(mx)] \\ & - \frac{i\tau_n}{mx} \left[m [x j_n(x)]' - a_n m [x h_n^{(1)}(x)]' - d_n [m x j_n(mx)]' \right] = 0, \end{aligned} \quad (2.17a)$$

$$\begin{aligned} \hat{\mathbf{e}}_\phi : \quad & \tau_n [j_n(x) - b_n h_n^{(1)}(x) - c_n j_n(mx)] \\ & - \frac{i\pi_n}{mx} \left[m [x j_n(x)]' - a_n m [x h_n^{(1)}(x)]' - d_n [m x j_n(mx)]' \right] = 0. \end{aligned} \quad (2.17b)$$

The primes represent differentiation with respect to the argument in parentheses. From here, we can obtain:

$$j_n(x) = b_n h_n^{(1)}(x) + c_n j_n(mx), \quad (2.18a)$$

$$m [x j_n(x)]' = a_n m [x h_n^{(1)}(x)]' + d_n [m x j_n(mx)]', \quad (2.18b)$$

and similarly from the magnetic field conditions:

$$\mu_1 j_n(x) = \mu m d_n j_n(mx) + \mu_1 a_n h_n^{(1)}, \quad (2.18c)$$

$$\mu_1 [x j_n(x)]' = b_n \mu_1 [x h_n^{(1)}(x)]' + \mu c_n [m x j_n(mx)]'. \quad (2.18d)$$

Finally, solving the above simultaneous equations, we arrive with scattering coefficients, a_n and b_n :

$$a_n = \frac{\mu m^2 j_n(mx) [x j_n(x)]' - \mu_1 j_n(x) [m x j_n(mx)]'}{\mu m^2 j_n(mx) [x h_n^{(1)}(x)]' - \mu_1 h_n^{(1)}(x) [m x j_n(mx)]'}, \quad (2.19a)$$

$$b_n = \frac{\mu_1 j_n(mx) [x j_n(x)]' - \mu j_n(x) [m x j_n(mx)]'}{\mu_1 j_n(mx) [x h_n^{(1)}(x)]' - \mu h_n^{(1)}(x) [m x j_n(mx)]'}, \quad (2.19b)$$

which can be simplified by using Riccati-Bessel functions, $\psi_n(\rho) = \rho j_n(\rho)$ and $\xi_n(\rho) = \rho h_n^{(1)}(\rho)$, and the assumption that the permeability of both mediums is the same.

$$a_n = \frac{m \psi_n(mx) \psi_n'(x) - \psi_n(x) \psi_n'(mx)}{m \psi_n(mx) \xi_n'(x) - \xi_n(x) \psi_n'(mx)}, \quad (2.20a)$$

$$b_n = \frac{\psi_n(mx) \psi_n'(x) - m \psi_n(x) \psi_n'(mx)}{\psi_n(mx) \xi_n'(x) - m \xi_n(x) \psi_n'(mx)}. \quad (2.20b)$$

2.1.5 The amplitude matrix, efficiency factors and Stokes parameters

The scattering and extinction efficiency factors, Q^{sca} and Q^{ext} , can be calculated directly from the amplitude coefficients [Grainger et al. 2004]:

$$Q^{\text{ext}} = \frac{2}{x^2} \sum_{n=1}^{\infty} (2n+1) \Re(a_n + b_n), \quad (2.21)$$

$$Q^{\text{sca}} = \frac{2}{x^2} \sum_{n=1}^{\infty} (2n+1) (|a_n|^2 + |b_n|^2). \quad (2.22)$$

The amplitude matrix, \mathbf{S} , is diagonal due to symmetry and takes the form

$$\begin{pmatrix} E_{\parallel s} \\ E_{\perp s} \end{pmatrix} = \frac{e^{ikr}}{kr} \begin{pmatrix} S_2 & 0 \\ 0 & S_1 \end{pmatrix} \begin{pmatrix} E_{\parallel i} \\ E_{\perp i} \end{pmatrix}, \quad (2.23)$$

where

$$S_1 = \sum_{n=1}^{\infty} \frac{2n+1}{n(n+1)} (a_n \pi_n + b_n \tau_n), \quad (2.24a)$$

$$S_2 = \sum_{n=1}^{\infty} \frac{2n+1}{n(n+1)} (a_n \tau_n + b_n \pi_n). \quad (2.24b)$$

Stokes parameters from scattered light can be calculated from the incident light's parameters using

the phase matrix, \mathbf{Z} , and the amplitude matrix:

$$\begin{pmatrix} I_s \\ Q_s \\ U_s \\ V_s \end{pmatrix} = \frac{1}{k^2 R^2} \begin{pmatrix} Z_{11} & Z_{12} & 0 & 0 \\ Z_{12} & Z_{11} & 0 & 0 \\ 0 & 0 & Z_{33} & Z_{34} \\ 0 & 0 & -Z_{34} & Z_{33} \end{pmatrix} \begin{pmatrix} I_i \\ Q_i \\ U_i \\ V_i \end{pmatrix}, \quad (2.25)$$

where

$$\begin{aligned} Z_{11} &= \frac{1}{2} (|S_2|^2 + |S_1|^2), & Z_{12} &= \frac{1}{2} (|S_2|^2 - |S_1|^2), \\ Z_{33} &= \frac{1}{2} (S_2^* S_1 + S_2 S_1^*), & Z_{34} &= \frac{i}{2} (S_2^* S_1 - S_2 S_1^*). \end{aligned}$$

2.2 T-matrix theory

T-matrix theory is a powerful method of formulating an analytical solution to EM scattering by non-spherical particles. The original theory was developed by P. C. Waterman [Waterman 1965, Waterman 1971] and is one of the most popular methods available, with many implementations for different uses [Mishchenko et al. 2008]. The power of the method lies in relating the expansion of the scattered waves' expansion coefficients in matrix notation to the incident coefficients. Changing the incident beam's angle, or the measurement angle of the scattered beam, does not require recalculation of the T-matrix. The methods of obtaining the matrix itself are varied.

2.2.1 Vector Spherical Wave functions

In the spherical case, symmetry and a careful choice of incoming light direction meant we could make the simplification that only the expansion coefficients where $m = 1$ in (2.13) and (2.14) were non-zero. This allowed us to write the scattering coefficients with only one subscript, n as in (2.20).

In a more general example, this is not the case and we have to deal with another dimension of scattering coefficients. For this section, the definition of the VSWF changes slightly, so as to fit that used by a large number of people in the T-matrix community. Gone are the even and odd versions of \mathbf{M} and \mathbf{N} and in their place are single functions containing $e^{im\phi}$. From (3) of Ch. 6 of Mishchenko

et al. [2000], we define (with notation and form altered to fit the conventions of this report):

$$\mathbf{M}_{mn}(k\mathbf{r}) = \gamma_{nm} z_n(kr) e^{im\phi} \left[\frac{im}{\sin\theta} P_n^m \hat{\mathbf{e}}_\theta - \frac{dP_n^m}{d\theta} \hat{\mathbf{e}}_\phi \right], \quad (2.26a)$$

$$\mathbf{N}_{mn}(k\mathbf{r}) = \gamma_{nm} \frac{e^{im\phi}}{kr} \left\{ n(n+1) z_n(kr) P_n^m \hat{\mathbf{e}}_r + \frac{d}{d(kr)} [kr z_n(kr)] \times \left[\frac{dP_n^m}{d\theta} \hat{\mathbf{e}}_\theta + \frac{im}{\sin\theta} P_n^m \hat{\mathbf{e}}_\phi \right] \right\}, \quad (2.26b)$$

which bear a striking resemblance to equations (A.1). The factor γ_{nm} is

$$\gamma_{nm} = \sqrt{\frac{(2n+1)(n-m)!}{4\pi n(n+1)(n+m)!}}. \quad (2.27)$$

Equations (2.26) are more generally written in terms of $\mathbf{B}_{mn}(\theta)$ and $\mathbf{C}_{mn}(\theta)$ which play a similar role to the π_n and τ_n in §2.1.4, being defined:

$$\mathbf{B}_{mn}(\theta) = (-1)^m \left(\frac{(n-m)!}{(n+m)!} \right)^{\frac{1}{2}} \times \left[\frac{dP_n^m}{d\theta} \hat{\mathbf{e}}_\theta + \frac{im}{\sin\theta} P_n^m \hat{\mathbf{e}}_\phi \right], \quad (2.28a)$$

$$\mathbf{C}_{mn}(\theta) = (-1)^m \left(\frac{(n-m)!}{(n+m)!} \right)^{\frac{1}{2}} \times \left[\frac{im}{\sin\theta} P_n^m \hat{\mathbf{e}}_\theta - \frac{dP_n^m}{d\theta} \hat{\mathbf{e}}_\phi \right], \quad (2.28b)$$

and immediately from (2.26), we have:

$$\mathbf{M}_{mn} = (-1)^m d_n z_n \mathbf{C}_{mn} e^{im\phi}, \quad (2.29a)$$

$$\mathbf{N}_{mn} = \frac{d_n}{kr} \times \left[n(n+1) z_n \left(\frac{(n-m)!}{(n+m)!} \right)^{\frac{1}{2}} \frac{P_n^m}{r} \hat{\mathbf{e}}_r + (-1)^m [kr z_n]' \mathbf{B}_{mn} \right] e^{im\phi}, \quad (2.29b)$$

where

$$d_n = \left(\frac{2n+1}{4\pi n(n+1)} \right)^{\frac{1}{2}}.$$

Now that there are new forms of the VSWF, we had better rewrite the incident and scattered fields:

$$\mathbf{E}_i(\mathbf{r}) = \sum_{n=1}^{\infty} \sum_{m=-n}^n [a_{mn} \mathbf{M}_{mn}^{(1)}(k\mathbf{r}) + b_{mn} \mathbf{N}_{mn}^{(1)}(k\mathbf{r})], \quad (2.30a)$$

$$\mathbf{E}_s(\mathbf{r}) = \sum_{n=1}^{\infty} \sum_{m=-n}^n [p_{mn} \mathbf{M}_{mn}^{(3)}(k\mathbf{r}) + q_{mn} \mathbf{N}_{mn}^{(3)}(k\mathbf{r})] \quad r > R_>. \quad (2.30b)$$

The scattered field is only correct outside of the sphere bounded by $R_>$, the smallest sphere which completely encompasses our shape as shown in Fig. 2.1. As previously in §2.1.2, the superscripts ⁽¹⁾ and ⁽³⁾ refer to the type of Bessel functions used in the radial solution.

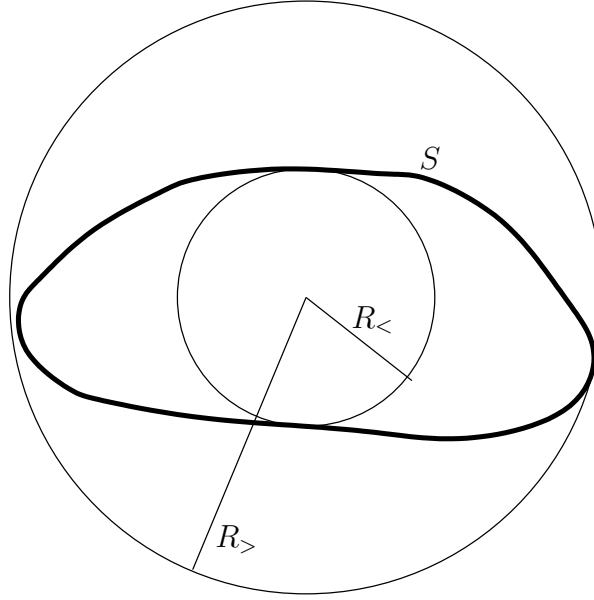


Figure 2.1: Scattering object bounded by surface S . $R_>$ is the radius of the smallest sphere which completely encloses S . $R_<$ is the largest sphere which can entirely fit inside S . Adapted from Mishchenko et al. [2000].

Since Maxwell's equations are linear, there must be a linear relationship between incident and scattered field coefficients. This we write as:

$$p_{mn} = \sum_{n'=1}^{\infty} \sum_{m'=-n'}^{n'} [T_{mnm'n'}^{11} a_{m'n'} + T_{mnm'n'}^{12} b_{m'n'}], \quad (2.31)$$

$$q_{mn} = \sum_{n'=1}^{\infty} \sum_{m'=-n'}^{n'} [T_{mnm'n'}^{21} a_{m'n'} + T_{mnm'n'}^{22} b_{m'n'}], \quad (2.32)$$

or in matrix notation,

$$\begin{bmatrix} \mathbf{p} \\ \mathbf{q} \end{bmatrix} = \begin{bmatrix} \mathbf{T}^{11} & \mathbf{T}^{12} \\ \mathbf{T}^{21} & \mathbf{T}^{22} \end{bmatrix} \begin{bmatrix} \mathbf{a} \\ \mathbf{b} \end{bmatrix} = \mathbf{T} \begin{bmatrix} \mathbf{a} \\ \mathbf{b} \end{bmatrix}. \quad (2.33)$$

So knowing the incident field and the T-matrix allows us to calculate the scattered field (see §2.2.2).

The incident electric field is written as

$$\mathbf{E}_i = \mathbf{E}_0 e^{i\mathbf{k}_i \cdot \mathbf{r}} = [E_\theta \hat{\mathbf{e}}_\theta(\theta_i, \phi_i) + E_\phi \hat{\mathbf{e}}_\phi(\theta_i, \phi_i)] e^{i\mathbf{k}_i \cdot \mathbf{r}}, \quad (2.34)$$

where \mathbf{E}_0 gives the amplitude and polarisation of the beam and the direction is given by $\mathbf{k}_i = k_i \hat{\mathbf{n}}_i$. The field is required in the form given by (2.30a) which once again means we have to follow a similar method to (2.10). The form of the incident coefficients are [Mishchenko et al. 2000, Tsang et al. 1985]:

$$a_{mn} = 4\pi(-1)^m i^n d_n \mathbf{C}_{mn}^*(\theta_i) \cdot \mathbf{E}_0 e^{im\phi_i}, \quad (2.35a)$$

$$b_{mn} = 4\pi(-1)^m i^{n-1} d_n \mathbf{B}_{mn}^*(\theta_i) \cdot \mathbf{E}_0 e^{im\phi_i}. \quad (2.35b)$$

2.2.2 The amplitude matrix and other important physical properties

Once the T-matrix has been obtained, we can calculate any properties we would sensibly want with relative ease. The amplitude matrix is written:

$$\begin{aligned} \mathbf{S}(\mathbf{n}_s, \mathbf{n}_i) = & \frac{4\pi}{k} \sum_{mnm'n'} i^{n'-n-1} (-1)^{m+m'} d_n d_{n'} e^{i(m\phi_s - m'\phi_i)} \\ & \times \left\{ [T_{mnm'n'}^{11} \mathbf{C}_{mn}(\theta_s) + T_{mnm'n'}^{21} i\mathbf{B}_{mn}(\theta_s)] \mathbf{C}_{m'n'}^*(\theta_i) \right. \\ & \left. - i [T_{mnm'n'}^{12} \mathbf{C}_{mn}(\theta_s) + T_{mnm'n'}^{22} i\mathbf{B}_{mn}(\theta_s)] \mathbf{B}_{m'n'}^*(\theta_i) \right\}. \quad (2.36) \end{aligned}$$

Note that for any specific shape, the T-matrix need be computed only once. The direction of incident and scattered light, \mathbf{n}_i and \mathbf{n}_s , then determines the value of \mathbf{S} [Mishchenko 2000].

2.2.3 Orientation averaging

In some cases, it is not necessary to look at scattering functions to calculate the orientation average of a scattering characteristic. The T-matrix itself will have this information. There is an obvious speed benefit here.

The orientation averaged T-matrix gives the scattering relationships for a particle, averaged over all orientations. From [Mishchenko et al. 2000], it is given as:

$$\langle T_{mnm'n'}^{ij} \rangle = \frac{1}{2n+1} \delta_{mm'} \delta_{nn'} \sum_{m_1=-n}^n T_{m_1nm_1n}^{ij}, \quad i, j = 1, 2. \quad (2.37)$$

The averaged cross sections are given by [Mishchenko 1990, Mishchenko et al. 1996]:

$$\langle C_{\text{ext}} \rangle = -\frac{2\pi}{k^2} \Re \sum_{n=1}^{\infty} \sum_{m=-n}^n [T_{mnmn}^{11} + T_{mnmn}^{22}], \quad (2.38)$$

$$\langle C_{\text{sca}} \rangle = \frac{2\pi}{k^2} \sum_{n=1}^{\infty} \sum_{n'=1}^{\infty} \sum_{m=-n}^n \sum_{m'=-n'}^{n'} \sum_{i=1}^2 \sum_{j=1}^2 |T_{mnm'n'}^{ij}|^2, \quad (2.39)$$

where the sums over n will be truncated at some appropriate value n_{max} .

2.2.4 Calculating the \mathbf{T} -matrix

Now that we have the T-matrix formalisation, we need to obtain the values of $T_{mnm'n'}^{ij}$. In this chapter, I shall only mention the one method of obtaining \mathbf{T} , but other tactics are available, some of which I intend to investigate in the next year (see Chapter 4).

The extended boundary condition method

The original tactic, devised by Waterman, was based on the Huygen's principle [Waterman 1969]. Weber's analogue of Helmholtz's theorem [Baker & Copson 1987], can be written as:

$$\text{At a point } \mathbf{r}, \quad \iint_S dS \hat{\mathbf{n}} \cdot [w \nabla H_0^{(1)} - H_0^{(1)} \nabla w] = \begin{cases} 0 & \text{if } \mathbf{r} \text{ lies within } S, \\ 4i w(\mathbf{r}) & \text{if } \mathbf{r} \text{ lies outside } S, \end{cases} \quad (2.40)$$

where $w(\mathbf{r})$ satisfies $(\nabla^2 + k^2)w = 0$ and $H_0^{(1)}$ are outgoing spherical Hankel functions. Physically, given sources inside a surface, S , with \mathbf{r} lying outside the surface, the total effect of these sources at \mathbf{r} is identical to a new set of sources covering the surface at all points dS . These can be seen as secondary waves.

Analogous to (2.10), incident fields can be calculated in terms of VSWFs. We can no longer take the liberty of assuming that incident light will approach our object along a line of rotational symmetry, but this just means there will be more coefficients in our expansion. From Morse & Feshbach [1953], we are given the orthogonality relations for the functions $\mathbf{M}_{mn}^{(1)}$, $\mathbf{N}_{mn}^{(1)}$, $\mathbf{M}_{mn}^{(3)}$ and $\mathbf{N}_{mn}^{(3)}$ over the integral of a closed surface.

$$\iint_S dS \hat{\mathbf{n}} \cdot [(\nabla \times \boldsymbol{\psi}_{mn}^{(1)}) \times \boldsymbol{\psi}'_{-m'n'}^{(3)} + \boldsymbol{\psi}_{mn}^{(1)} \times (\nabla \times \boldsymbol{\psi}'_{-m'n'}^{(3)})] = \frac{(-1)^m}{ik} \delta_{mm'} \delta_{nn'} \delta_{\psi\psi'}, \quad (2.41)$$

where $\boldsymbol{\psi}$ is \mathbf{M} or \mathbf{N} . This can be shown by integrating over a sphere with vanishing radius for $\boldsymbol{\psi}^{(1)}$,

where $j_n(kr) \rightarrow 0$ for $kr \rightarrow 0$ and integrating over a sphere with $r \rightarrow \infty$ for $\psi^{(3)}$, where $h_n^{(1)}(kr) \rightarrow 0$.

Equation (2.41) can be used as a test function for finding the expansion coefficients for incident, internal and scattered fields. The closed surface over which we integrate is unimportant, and the solution is the same as we stretch this surface, for example, into a sphere, for ease of calculations, or the shape of a particle, for correct physical conditions.

Outside of the scattering particle, the total electric field can be written as $\mathbf{E} = \mathbf{E}_i + \mathbf{E}_s$. The forms of \mathbf{E}_i and \mathbf{E}_s given in (2.30) have orthogonal VSWF, so that, for example:

$$\begin{aligned} \iint_S dS \hat{\mathbf{n}} \cdot [(\nabla \times \mathbf{E}_i) \times \mathbf{M}_{-mn}^{(3)} + \mathbf{E}_i \times (\nabla \times \mathbf{M}_{-mn}^{(3)})] \\ = \iint_S dS \hat{\mathbf{n}} \cdot [(\nabla \times \mathbf{E}) \times \mathbf{M}_{-mn}^{(3)} + \mathbf{E} \times (\nabla \times \mathbf{M}_{-mn}^{(3)})] = \frac{(-1)^m}{i k} a_{mn}, \end{aligned} \quad (2.42)$$

and similarly for the other coefficients, b_{mn} , p_{mn} and q_{mn} . Continuing with a_{mn} and using the identities (2.1) and (2.4) which still hold for the new definitions of \mathbf{M}_{mn} and \mathbf{N}_{mn} , we use (2.42) and say:

$$a_{mn} = i k (-1)^m \iint_S dS \hat{\mathbf{n}} \cdot [(\nabla \times \mathbf{E}) \times \mathbf{M}_{-mn}^{(3)} + \mathbf{E} \times (\nabla \times \mathbf{M}_{-mn}^{(3)})] \quad (2.43)$$

$$= i k (-1)^m \iint_S dS \hat{\mathbf{n}} \cdot [i\omega\mu \mathbf{H} \times \mathbf{M}_{-mn}^{(3)} + k\mathbf{E} \times \mathbf{N}_{-mn}^{(3)}] \quad (2.44)$$

and using the well known vector identity $\mathbf{a} \cdot (\mathbf{b} \times \mathbf{c}) = (\mathbf{a} \times \mathbf{b}) \cdot \mathbf{c}$,

$$= i k (-1)^m \iint_S dS [i\omega\mu (\hat{\mathbf{n}} \times \mathbf{H}) \cdot \mathbf{M}_{-mn}^{(3)} + k(\hat{\mathbf{n}} \times \mathbf{E}) \cdot \mathbf{N}_{-mn}^{(3)}]. \quad (2.45)$$

Thus we arrive at the equations,

$$\begin{pmatrix} a_{mn} \\ b_{mn} \end{pmatrix} = i k (-1)^m \iint_S dS \left[i\omega\mu (\hat{\mathbf{n}} \times \mathbf{H}) \cdot \begin{pmatrix} \mathbf{M}_{-mn}^{(3)} \\ \mathbf{N}_{-mn}^{(3)} \end{pmatrix} + k(\hat{\mathbf{n}} \times \mathbf{E}) \cdot \begin{pmatrix} \mathbf{N}_{-mn}^{(3)} \\ \mathbf{M}_{-mn}^{(3)} \end{pmatrix} \right], \quad (2.46a)$$

$$\begin{pmatrix} p_{mn} \\ q_{mn} \end{pmatrix} = i k (-1)^m \iint_S dS \left[i\omega\mu (\hat{\mathbf{n}} \times \mathbf{H}) \cdot \begin{pmatrix} \mathbf{M}_{-mn}^{(1)} \\ \mathbf{N}_{-mn}^{(1)} \end{pmatrix} + k(\hat{\mathbf{n}} \times \mathbf{E}) \cdot \begin{pmatrix} \mathbf{N}_{-mn}^{(1)} \\ \mathbf{M}_{-mn}^{(1)} \end{pmatrix} \right]. \quad (2.46b)$$

On the surface of our scatterer, we can expand the surface fields in terms of VSWF with wavenumber $k_s = m k$. The internal field should match the sum of incident and scattered fields at the boundary, so our surface fields must be made of regular spherical functions, $\mathbf{M}_{mn}^{(1)}$ and $\mathbf{N}_{mn}^{(1)}$, since the internal field

cannot have singularities at the origin. We can write the electric surface fields as

$$\hat{\mathbf{n}} \times \mathbf{E} = \hat{\mathbf{n}} \times \sum_{mn} [c_{mn} \mathbf{M}_{mn}^{(1)}(mk\mathbf{r}) + d_{mn} \mathbf{N}_{mn}^{(1)}(mk\mathbf{r})], \quad (2.47a)$$

and again using (2.1) and (2.4) we can get the magnetic surface field,

$$\hat{\mathbf{n}} \times \mathbf{H} = \hat{\mathbf{n}} \times \sum_{mn} \frac{mk}{i\omega\mu} [d_{mn} \mathbf{M}_{mn}^{(1)}(mk\mathbf{r}) + c_{mn} \mathbf{N}_{mn}^{(1)}(mk\mathbf{r})]. \quad (2.47b)$$

Now, substituting (2.47) into (2.46) we can relate both the incident and scattering coefficients to the interior field coefficients. In matrix form these relationships can be written as

$$\begin{bmatrix} \mathbf{a} \\ \mathbf{b} \end{bmatrix} = \mathbf{Q}_1 \begin{bmatrix} \mathbf{c} \\ \mathbf{d} \end{bmatrix}, \quad \begin{bmatrix} \mathbf{p} \\ \mathbf{q} \end{bmatrix} = \mathbf{Q}_2 \begin{bmatrix} \mathbf{c} \\ \mathbf{d} \end{bmatrix}.$$

Finally, we can invert \mathbf{Q}_1 and obtain the scattered field from the incident field by:

$$\begin{bmatrix} \mathbf{p} \\ \mathbf{q} \end{bmatrix} = \mathbf{Q}_2 \mathbf{Q}_1^{-1} \begin{bmatrix} \mathbf{a} \\ \mathbf{b} \end{bmatrix}, \quad (2.48)$$

where we have now discovered the T-matrix,

$$\mathbf{T} = \mathbf{Q}_2 \mathbf{Q}_1^{-1}. \quad (2.49)$$

All of the integrals used to calculate \mathbf{Q}_1 and \mathbf{Q}_2 are easily obtained. The integrands (after constant terms have been taken outside the integral) are the products of VSWFs and so are known. The integral is carried out over the surface of the scattering object which can be defined as $r = r(\theta, \phi)$. We can write $dS \hat{\mathbf{n}}$ as

$$dS \hat{\mathbf{n}} = r^2 \sin \theta \left[\hat{\mathbf{e}}_r - \frac{1}{r} \frac{\partial r}{\partial \theta} \hat{\mathbf{e}}_\theta - \frac{1}{r \sin \theta} \frac{\partial r}{\partial \phi} \hat{\mathbf{e}}_\phi \right] d\theta d\phi. \quad (2.50)$$

The time consuming (and numerically tricky) part of the above operation to obtain \mathbf{T} is the inversion of \mathbf{Q}_1 , whose elements can vary by several orders of magnitude and which can be a very large matrix. The inversion process can, in many cases, be ill conditioned.

Tsang et al. [1985] relied on Green's dyadic in their method of calculation, but using Waterman [1979] it was possible to avoid this; an advantage, since the use of Green's dyadic requires fictitious fields on the interior of the scattering object, which muddy the physical picture.

2.3 Fractal representation

A fractal is a rough or fragmented geometric shape that can be split into parts, each of which is (at least approximately) a reduced-size copy of the whole. For this reason fractals are said to be scale invariant (see Fig. 2.2). Aggregates of aerosols can be thought of as fractal particles. As small aerosols

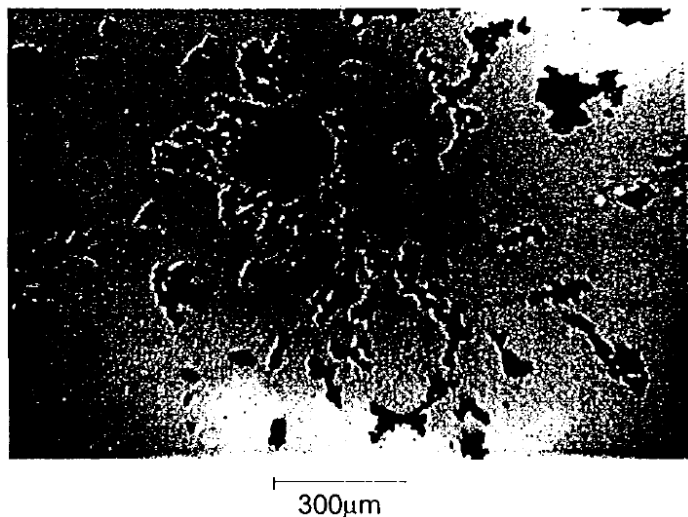
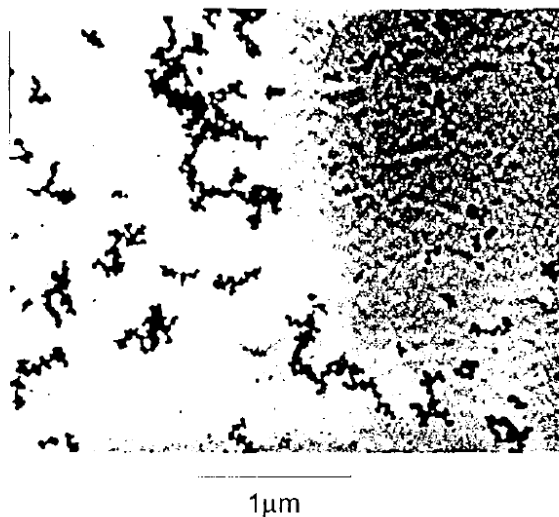


Figure 2.2: Soot fractal aggregates. The upper picture was taken with a TEM and the lower with an optical microscope with the scale two and a half orders of magnitude less. The same shapes can be seen at both length scales, showing the size invariant nature of fractal aggregates. This is Fig. 1 in Sorensen [2001]

collide and join, they become bigger and then can collide with other small aggregate particles to form larger aggregates etc.

The method (outlined by Sorensen [2001]) assumes that individual molecules in the aggregate are small spheres in the Rayleigh limit ($|m|kr_0 \ll 1$) so that Mie theory need not be applied. Aggregates are given a single parameter, the fractal dimension D , which relates the number of monomers in an

aggregate to the total size of the aggregate in the equation:

$$N \propto \left(\frac{R}{r_0}\right)^D, \quad (2.51)$$

where N is the number of monomers in the aggregate and will be proportional to the mass, R is a measure of the overall aggregate radius (the radius of gyration) and r_0 is the radius of an individual spherical monomer. The value of the fractal dimension must be less than the number of dimensions in which the scatterer is represented, d .

2.3.1 The scaling approach

As a lowest order approach, we can look at scaling, which gives a nice clear picture of the ideas without cumbersome theory. Defining $q = 4\pi\lambda^{-1}\sin(\theta/2)$, where θ is the scattering angle, we split our scattering object into “q-regions” of radius q^{-1} (see Fig. 2.3) where it is assumed that within “q-regions”, the scattered light is in phase at the detector. Thus the scattered intensity from a single

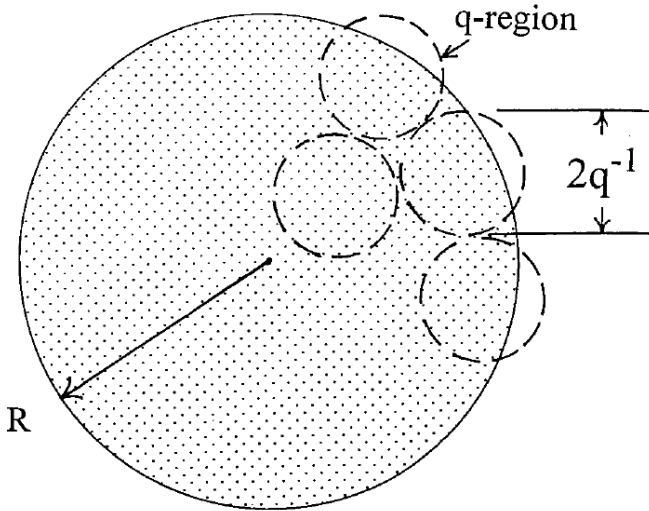


Figure 2.3: A system of point scatterers in a region of size R showing the definition of q-regions. This is Fig. 4 in Sorensen [2001].

q-region is $I_q \propto N_q^2$ where N_q is the number of scatterers in the region. The Ewald-Oseen Extinction Theorem [Hecht 2002, p. 95] say that within an interior of uniform density, primary waves are precisely cancelled out by scattered term. This means that only q-regions on the surface of the system contribute to the light scattering. The scattered waves from separate q-regions add randomly so that the number of q-regions on the surface, n_q , is proportional to the intensity measured and we can say:

$$I(q) \propto n_q N_q^2. \quad (2.52)$$

There are three regimes of q .

In the case where $q < R_g^{-1}$ we can see that $n_q = 1$, $N_q = N$ and so (2.52) gives us $I(q) \propto N^2$. Physically, the assumption that all scatterers in the system are in phase means that the intensity is nine times that for a single scatterer.

In the case where $R_g^{-1} < q < r_0^{-1}$, we need to use the fractal dimension. The number of scatterers in a q -region is proportional to the total number of scatterers, N , and the ratio of the length scale, R_g , to the q length scale, q^{-1} to the power of the fractal dimension (since the mass will also have a fractal dimension and N is proportional to mass), $N_q = N(q^{-1}/R)^D$. Similarly, the number of q -regions on the surface, n_q , is proportional to the ratio of the two characteristic lengths, $n_q \propto (R/q)^D$ so that in this case, $I(q) \propto N^2(qR)^{-2D}(qR)^D = N^2(qR)^{-D}$.

When $q > r_0$, all the scatterers have random phases with others so $I(q)$ is proportional to the number of scatterers on the surface (due to Ewald-Oseen Extinction Theorem) and thus we have a final set of equations:

$$I(q) \propto \begin{cases} N^2 & \text{for } q < R^{-1}, \\ N^2(qR)^{-D} & \text{for } R^{-1} < q < r_0^{-1}, \\ N^2(R/r_0)^{-D} & \text{for } q > r_0^{-1}. \end{cases} \quad (2.53)$$

Full solutions now require that the point like scatterers used above are replaced by spheres. Methods of carrying out these calculations are varied and mentioned in [Sorensen 2001] which is a very useful review paper on fractal methods.

2.4 Examples of Scattering calculations

The following calculations demonstrate the added complexity of non-spherical particles to the phase function. Fig. 2.4 shows spherical phase functions calculated using Mie theory. Figures 2.5 and 2.6 demonstrate the variation in phase function with change in orientation of the particle, relative to the incident light direction. In the cases where the orientation of the particle is along a line of symmetry, we see that the phase function is symmetric about 0° .

Fig. 2.7 shows log normal distributions of spheroidal particles with various ratios of major and minor axes (a/b). Happily, because we are looking at the scattering from a uniform ensemble of particles, orientation averaging removes the asymmetry of the phase function. Notice that the main differences mainly occur in the back scatter direction which is unfortunate for nadir viewing satellites, if non-spherical particles are scattering light in the atmosphere.

The T-matrix calculations were obtained using the IDL DLM written to call Michael Mishchenko's

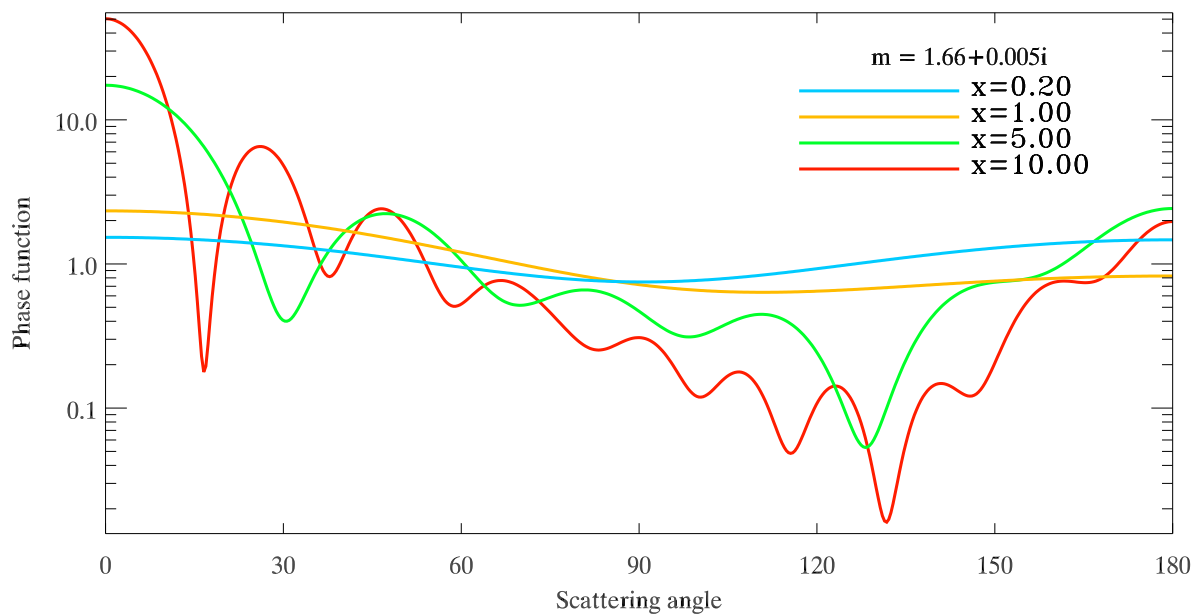


Figure 2.4: Example calculations of spherical phase functions. These were calculated using the EODG Mie scattering routines [*Light scattering routines* 2008]. The particles all have refractive index $m = 1.66 + 0.005i$. Particles with larger size parameters have more structure in their phase function.

FORTTRAN T-matrix code [*T-Matrix Codes for Computing Electromagnetic Scattering by Nonspherical and Aggregated Particles* 2008] in IDL.

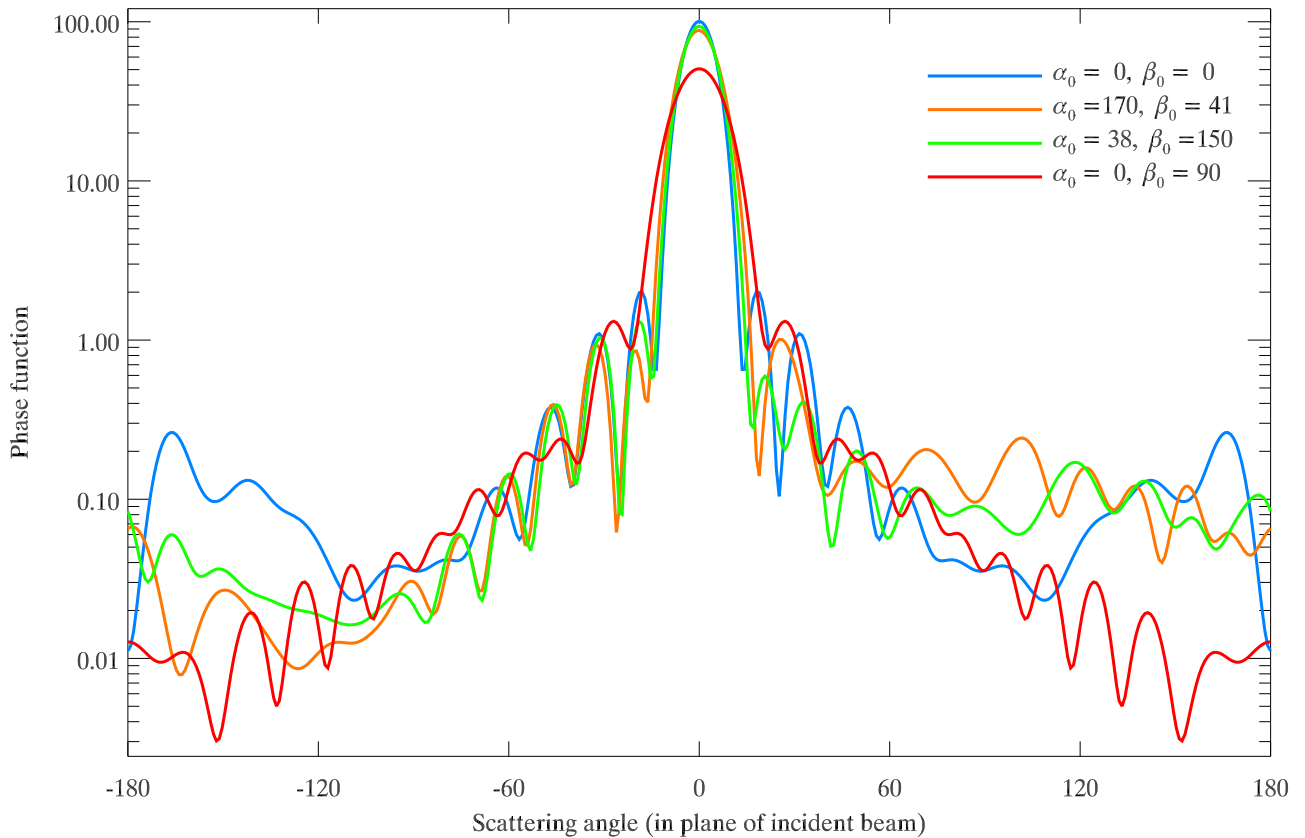


Figure 2.5: Comparing the scattering from an oblate spheroid in various fixed orientations. All of the lines are from the same particle with the following properties: $r = 1.1 \mu\text{m}$, $m = 1.9 + 0.05i$, $\lambda = 550 \text{ nm}$, $a/b = 1.8$. The radius, r , is the radius that a sphere would have in order to have the same surface area as the spheroid. Light enters along the z -axis and the scattering is in the y - z plane. The Euler angles α_0 and β_0 give the orientation of the spheroid relative to the incoming beam.

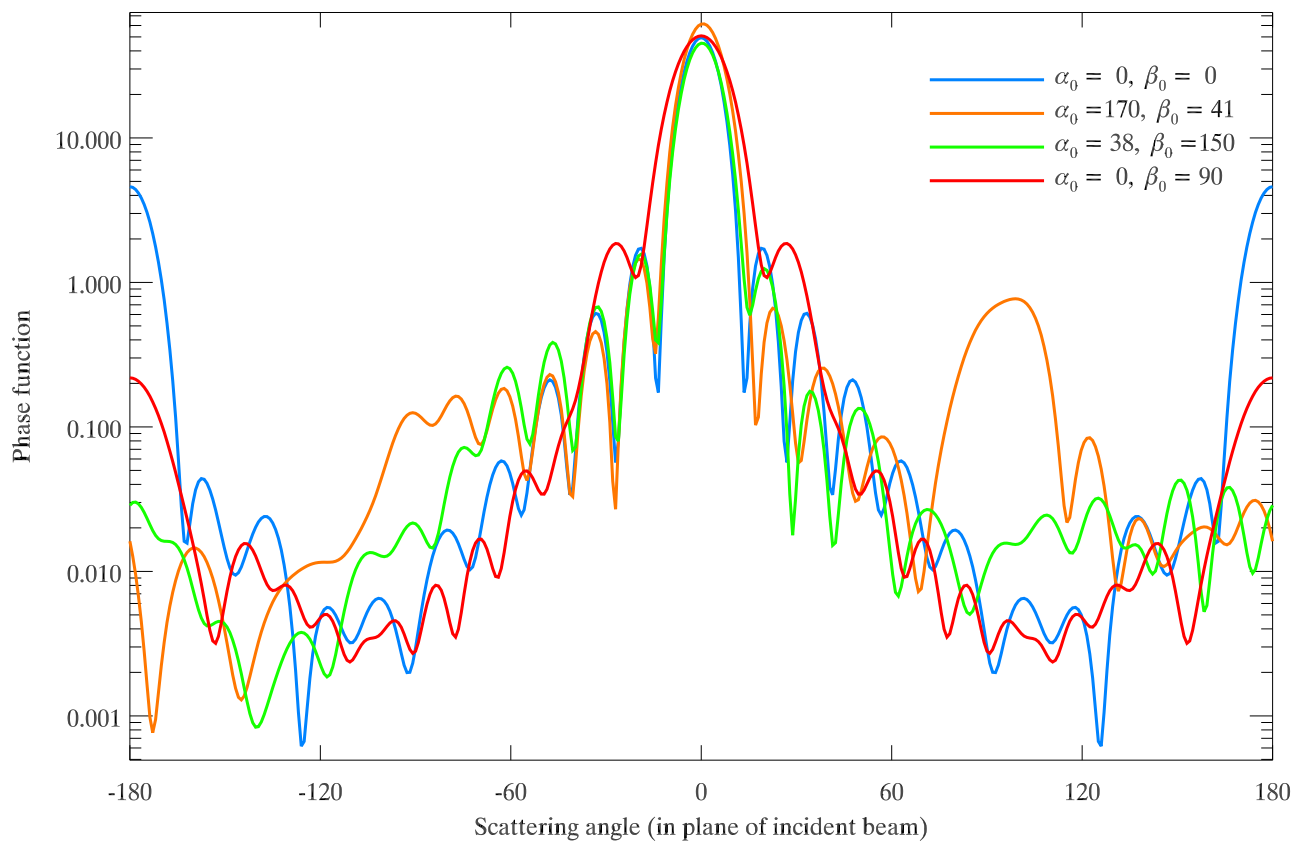


Figure 2.6: Comparing the scattering for an oblate cylinder with $d/l = 1.8$ and all other properties identical to Fig. 2.5.

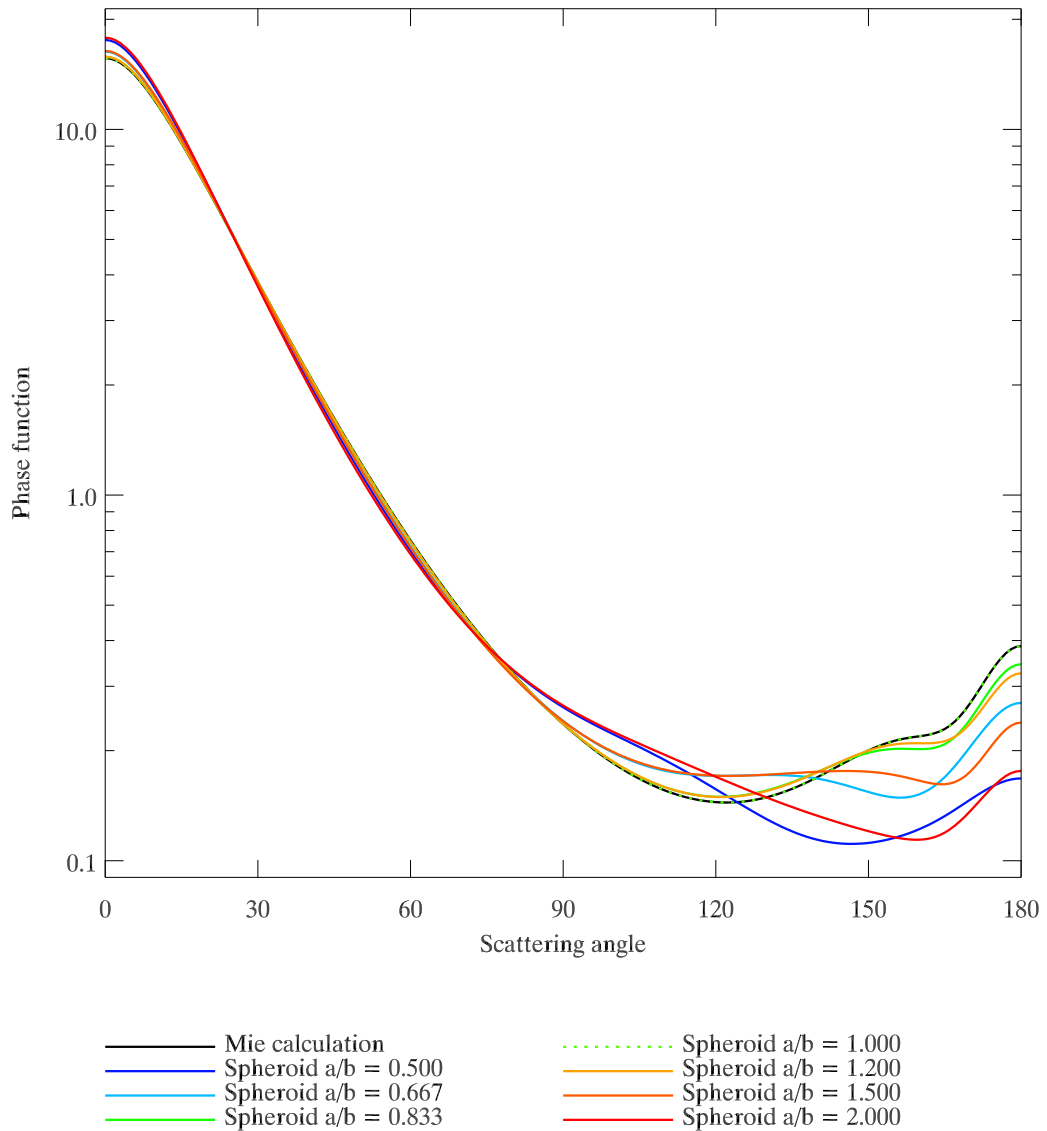


Figure 2.7: Comparing the scattering from log-normal distributions of different spheroids. For the above plots all distributions have identical characteristics, except for the aspect ratio of the particles. All distributions are log normal, with mean effective radius, $r_m = 0.2 \mu\text{m}$, refractive index, $m = 1.53 + 0.0055i$. The wavelength of scattered light is $\lambda = 863 \text{ nm}$ and the spread of the distribution is $S = 1.77$, where this means that the standard deviation of the natural log of mean radius, $\ln(r_m)$, is $\ln(S)$. Spheroid distributions are calculated using the IDL DLM written to run Mishchenko's T-matrix code [*T-Matrix Codes for Computing Electromagnetic Scattering by Nonspherical and Aggregated Particles* 2008]. The test case for a spheroid with $a/b = 1$ (a sphere) shows very good agreement with the Mie theory. Since this is a distribution of particles, the light scattering has been orientation averaged.

Chapter 3

SPARCLE Design and Optimisation

SPARCLE is a detector designed to measure the size and refractive index of particles. The setup is a capillary tube which pumps single particles through a cavity containing a laser beam and two light detectors. The laser's line of sight is across the path of passing particles so that light is scattered as they pass through the light beam. The light detectors will then be used to measure the phase function of the particle in the hopes of inferring its physical properties. The system is made such that the likelihood of more than one particle being in the cavity at once is prohibitively small.

The optimisation required was to find out where the most sensible place to put the two detectors is. In order to simplify the problem, the detectors were both placed in the same scattering plane¹.

3.1 The forward model

The forward model is our best representation of how the light scattered by a known particle will be measured by the system we define. The forward model is based on EODG Mie scattering routines [*Light scattering routines* 2008]. Since this code calculates the scattered light at a point, the detectors' measured intensity is calculated by summing the signal from a series of points.

Additional code² converts the light intensity which would arrive at a detector into the electrical current that would be measured. Flags alert the forward model to cases when the measurement will fail (for example, due to high signal to noise, or saturation of the detector).

Two different types of detector are used, a photo-multiplier tube (PMT) and a linear diode array (LDA). The PMT has a large surface area and is very sensitive to small changes in the amplitude of light. In comparison, the LDA is actually a large number of individual small detectors spaced close

¹The scattering plane is the plane that includes the incident and scattering light vectors.

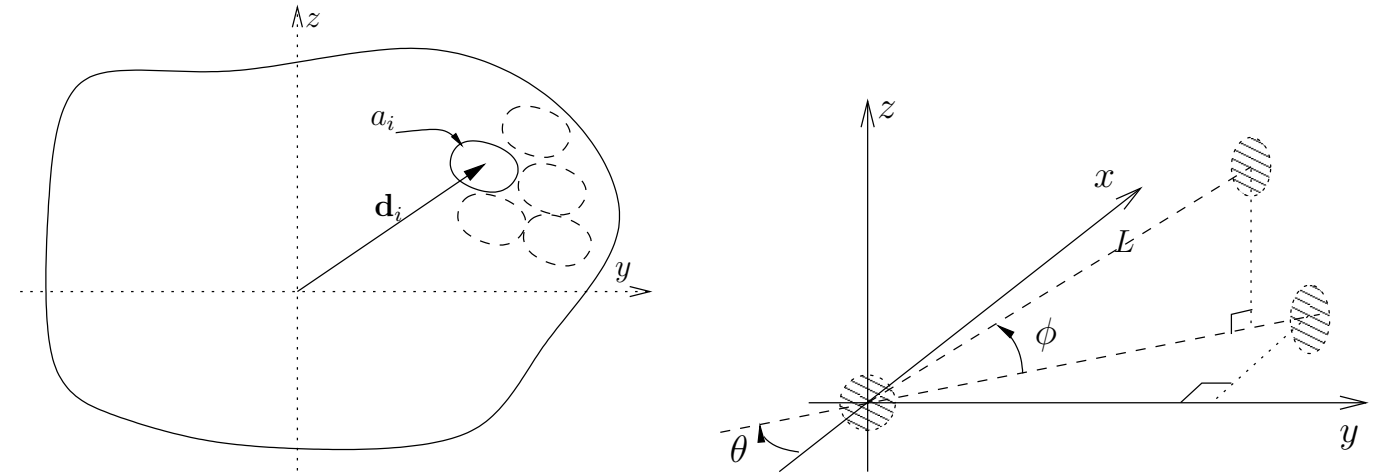
²Written by Dan Peters (dpeters@atm.ox.ac.uk).

together.

3.1.1 Detector geometry

Position

The incoming light travels along the x-axis. We assume that the detection areas of our PMT and LDA are in a 2D plane. The detector is defined as a series of points, \mathbf{d}_i with associated areas a_i . Initially, we define the detector elements in the y-z plane as shown in Fig. 3.1a. The whole detector can be defined by $\mathbf{D} = (\mathbf{d}_1, \mathbf{d}_2, \dots)$ and $\mathbf{A} = (a_1, a_2, \dots)$.



(a) Defining the detector in the y-z plane. Each element of detector, i , is given a position, \mathbf{d}_i , and an area, a_i .

(b) Placement of the detector - Once orientated, the detector is moved a distance L from the light scattering particle. It is then rotated an angle θ in the x-y plane, then ϕ out of the x-y plane.

Figure 3.1: Positioning the detector.

Now that we have a detector defined, we need to move it to a useful position. For this we employ matrix transformations. Initially it is at the origin, facing the incoming light. We can place it in any position we want using a translation and two rotations. If the detector is not to be orientated perpendicular to light scattering from the origin (defined as the location of the measured particle) then there must be a further three rotations.

Since matrix rotations map about the origin, any rotations carried out while the detector is placed there change its orientation relative to the path of the scattered beam. If the detector is not at the origin, the rotations move it to a new scattering angle, but maintain its orientation relative to the scattered beam.

The operations performed, for a detector starting at the origin, are thus:

1. Rotate an angle φ_1 about the x-axis.
2. Rotate an angle φ_2 about the y-axis.
3. Rotate an angle φ_3 about the z-axis.
4. Translate a distance L along the z-axis.
5. Rotate an angle ϕ about the y-axis.
6. Rotate an angle θ about the x-axis.

Positioning relative to the origin is shown in Fig. 3.1b. The final position of a detector element can be written as:

$$\mathbf{d}'_i = \mathbf{R}_\theta \mathbf{R}_\phi (\mathbf{R}_3 \mathbf{R}_2 \mathbf{R}_1 \mathbf{d}_i + \mathbf{t}), \quad (3.1)$$

where \mathbf{R} are rotation matrices and $\mathbf{t} = (L, 0, 0)$ is a translation. The angular position of an element, (θ_i, ϕ_i) , can be recovered by:

$$\cos \theta_i = \frac{x_i}{\sqrt{x_i^2 + y_i^2}}, \quad \text{and} \quad \cos \phi_i = \sqrt{\frac{x_i^2 + y_i^2}{x_i^2 + y_i^2 + z_i^2}}. \quad (3.2)$$

Surface area

For any flat detector surface with area a_i , the solid angle subtended from the origin to an element is,

$$d\Omega = \frac{\hat{\mathbf{n}} d\mathbf{A}}{r^2}, \quad (3.3)$$

where $\hat{\mathbf{n}}$ is the direction from the origin to that element, and $d\mathbf{A}$ is the element surface area in the direction normal to the plane. r is the distance from the origin to the element. Fig. 3.2 shows such a detector plane, S' , with the transformed origin ($\mathbf{0} \rightarrow \mathbf{0}'$) and an arbitrary transformed detector point, \mathbf{d}'_i . Using (3.3), we can see that in this case, with $\hat{\mathbf{n}} = \hat{\mathbf{d}}'_i$ and $r = |\mathbf{d}'_i|$ the transformed solid angle is given by:

$$d\Omega_i = \frac{\mathbf{d}'_i}{|\mathbf{d}'_i|^3} \cdot d\mathbf{S}'_i. \quad (3.4)$$

All that remains to be found now is $d\mathbf{S}'_i$ which will be the normal to the surface, S' , multiplied by the predefined area of the element, a_i .

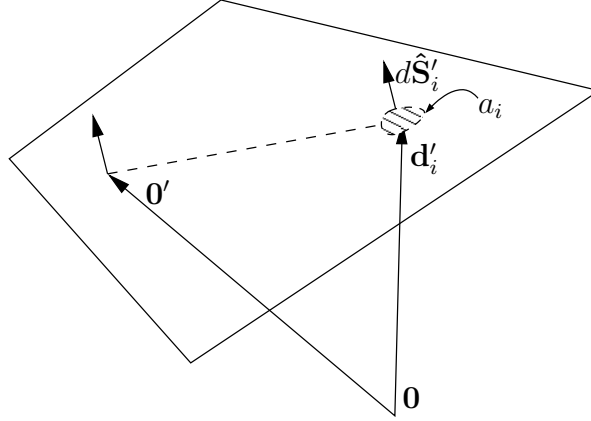


Figure 3.2: Showing the plane of the detector, S' , after being rotated and moved from the origin, $\mathbf{0}$.

We can define vectors, \mathbf{d}_0 , \mathbf{d}_y and \mathbf{d}_z , as

$$\mathbf{d}_0 = \mathbf{0} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}; \quad \mathbf{d}_y = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}; \quad \mathbf{d}_z = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}.$$

Taking the cross product of any two non-parallel vectors on the plane S' , will give us the surface normal, so we can say

$$d\mathbf{S}'_i = a_i \frac{(\mathbf{d}'_y - \mathbf{d}'_0) \wedge (\mathbf{d}'_z - \mathbf{d}'_0)}{|(\mathbf{d}'_y - \mathbf{d}'_0) \wedge (\mathbf{d}'_z - \mathbf{d}'_0)|},$$

and we find

$$d\Omega'_i = a_i \frac{\mathbf{d}'_i \cdot [(\mathbf{d}'_y - \mathbf{d}'_0) \wedge (\mathbf{d}'_z - \mathbf{d}'_0)]}{|\mathbf{d}'_i|^3 |(\mathbf{d}'_y - \mathbf{d}'_0) \wedge (\mathbf{d}'_z - \mathbf{d}'_0)|}. \quad (3.5)$$

3.1.2 The phase function and its derivatives

For a particle with state vector $\mathbf{x} = \{r, m_r, m_i\}$ and with incident light of wavelength λ polarised at an angle ξ to the perpendicular, the point like forward model defines irradiance, \mathbf{f} , and rate of change,

\mathbf{K}_x , at an angle θ in the plane of scattering as

$$\mathbf{f}(\mathbf{x}, \theta, \xi) = \frac{1}{4\pi} \frac{2}{x^2} [i_1(\mathbf{x}, \theta) \cdot \cos^2 \xi + i_2(\mathbf{x}, \theta) \cdot \sin^2 \xi], \quad (3.6)$$

$$\mathbf{K}_x(\mathbf{x}, \theta, \xi)|_r = \frac{1}{4\pi} \frac{2k}{x^2} \left[\frac{di_1}{dx} \cos^2 \xi + \frac{di_2}{dx} \sin^2 \xi - \frac{2}{x} (i_1 \cos^2 \xi + i_2 \sin^2 \xi) \right], \quad (3.7)$$

$$\mathbf{K}_x(\mathbf{x}, \theta, \xi)|_{m_r} = \frac{1}{4\pi} \frac{2}{x^2} \left[\frac{di_1}{dm_r} \cos^2 \xi + \frac{di_2}{dm_r} \sin^2 \xi \right], \quad (3.8)$$

$$\mathbf{K}_x(\mathbf{x}, \theta, \xi)|_{m_i} = \frac{1}{4\pi} \frac{2}{x^2} \left[\frac{di_1}{dm_i} \cos^2 \xi + \frac{di_2}{dm_i} \sin^2 \xi \right]. \quad (3.9)$$

The intensities of the two orthogonal polarisations are i_1 and i_2 , and x is the size parameter. Additional terms in (3.7) appear because the Mie scattering routines [*Light scattering routines* 2008] calculate the rate of change of intensity with respect to size parameter, not radius. For an entire detector, the total light measured, and its weighting function are defined:

$$\mathbf{f}_{\text{detector}} = \sum_i d\Omega_i \cdot \mathbf{f}(\mathbf{x}, \theta_i, \xi'_i), \quad (3.10)$$

$$\mathbf{K}_{\text{detector}} = \sum_i d\Omega_i \cdot \mathbf{K}_x(\mathbf{x}, \theta_i, \xi'_i). \quad (3.11)$$

The effective polarisation angle, ξ'_i , allows for pixels which lie outside the plane of $d_{iz} = 0$ and is defined

$$\xi'_i = \arcsin \left(\frac{d_{iz}}{\sqrt{d_{ix}^2 + d_{iz}^2}} \right) - \xi, \quad (3.12)$$

with ξ being defined as the angle of polarisation in the x-z plane from the z-axis.

Converting signal from arbitrary units to current across a detector

The incident photon flux (F_n) of the laser beam is obtained and the scattered photon radiance (photons.s⁻¹str⁻¹) measured is defined as [Peters 2007]

$$L_i = \frac{\sigma_{\text{scat}} F_n}{d\Omega_i} \mathbf{f}(\mathbf{x}, \theta_i, \xi'_i). \quad (3.13)$$

This is passed into further code which calculates the current that this would generate through the detector and the signal to noise ratio (which is important in §3.2) as well as flags for events in which reliable data would not be obtained due to measurement errors (eg PMT saturation or measurements outside the dynamic range of the detection electronics).

3.2 Optimisation method

In order to find the best position for the two detectors (the LDA and the PMT), a measure of effectiveness of a measurement should be defined. Two different values, the degrees of freedom and the information content can be used. Both of these values are calculated from the averaging kernel, defined in Rodgers [2000] as

$$\mathbf{A} = (\mathbf{S}_{\mathbf{x}_a}^{-1} + \mathbf{K}_x^T \mathbf{S}_y^{-1} \mathbf{K}_x)^{-1} \mathbf{K}_x^T \mathbf{S}_y^{-1} \mathbf{K}_x, \quad (3.14)$$

where $\mathbf{S}_{\mathbf{x}_a}$ is the covariance of a priori state and is known. The a priori values are an initial guess of how the system is expected to look. Covariance of the measurement vector, \mathbf{S}_y , is obtained from knowledge of the system. The averaging kernel, \mathbf{A} , gives the sensitivity of a retrieval to the true state.

Degrees of freedom

The number of degrees of freedom, d_s , gives a measure of the number of independent quantities that can be measured by a certain setup [Rodgers 2000]. Since SPARCLE is designed to retrieve particle radius, and the real and imaginary parts of refractive index, the range of values in this case is $0 \leq d_s \leq 3$.

$$d_s = \text{tr}(\mathbf{A}). \quad (3.15)$$

Information content

The Shannon information content, H , is a measure of the factor by which knowledge of a quantity would be improved by making a measurement. The logarithm of this factor is generally taken [Rodgers 2000].

$$H = -\frac{1}{2} \log_2 |\mathbf{I} - \mathbf{A}|, \quad (3.16)$$

where \mathbf{I} is the identity matrix. The larger the number the better.

Choosing the best measuring angles

In order to select the optimum detector positions, a range of radii and refractive indices are chosen which represent the full range of species hoped to be seen. The selected parameters are shown in Table 3.1.

The value of d_s and H is calculated for every single permutation of parameters. Then, for set distance and polarisation, averages of d_s and H are taken over all radii and refractive indices with all conditions given equal weight.

Parameter	Values
Radius (μm)	0.05, 0.1, 0.2, 0.5, 1, 2, 5
\Re (refractive index)	1.3 \rightarrow 2.4 in steps of 0.1
\Im (refractive index)	0.0 \rightarrow 0.5 in steps of 0.1
λ (μm)	0.5
Detection angle, θ (for both PMT and LDA)	$0^\circ \rightarrow 180^\circ$ in steps of 2°
Polarisation angle (from perpendicular)	0, $\pi/4$, $\pi/2$
Distance from source to detector (mm)	30, 60

Table 3.1: Parameters used in optimisation runs

This can be plotted as a contour with detection angle of the PMT and LDA as the two axes. The optimum angles, $\theta_{\text{PMT}}, \theta_{\text{LDA}}$ are easily obtained by finding the maximum values of d_s and H .

3.3 Conclusions

The results suggest that the best positioning of the detectors is with the PMT looking at the backscattered light, close to $\theta_{\text{PMT}} \simeq 180^\circ$. Unfortunately, this is unlikely to be possible since the laser is, by definition of backscatter, placed at 180° . It is also unfeasible to place a detector in the forward scatter direction, as the laser would saturate measurements at all times, and the sensors could become damaged. Because of this, optimum angles were recalculated with limits placed on the detector angles of $10^\circ \leq \theta_{\text{PMT}}, \theta_{\text{LDA}} \leq 140^\circ$. These still show that the PMT should be placed as far towards backscatter as possible. This is probably because the PMT is very sensitive to intensity, but has very poor spatial resolution (at 30mm, it has an angular range of almost 45°). Since the light in the forward direction is very intense, the PMT is likely to saturate in many cases, whereas the changes in the backscatter intensity are more smooth, and of lesser intensity.

By contrast, the LDA seems to prefer angles close to 30° . Physically, I think that the rapid changes in intensity and structure suit the instrument, which has 512 separate diodes, effectively measuring the structure of the phase function over 30° at high resolution.

One could argue that it makes sense that the maximum degrees of freedom are found when the detectors are placed further apart, and thus able to sample more of the phase function.

The closer the detector is to the source, the better. One could explain this by considering the signal to noise ratio, which is bound to be smaller when more light enters the detector.

Of the two plane polarisation modes, light polarised parallel to the plane of scattering ($\xi = 90^\circ$) would seem to have slightly more information, but this difference is not significant and, overall, I would

argue that the orientation of the beam's polarisation is unimportant.

Plots of the optimisation calculations are shown in Figs. 3.3, 3.4 and 3.5. d_s looks like a better measure of instrument reliability as the contours of constant d_s are far smoother than those of information content. That information content (H) was a less suitable method can be seen by comparing the upper and lower parts of Figures 3.3 and 3.4.

With the increasing size of the particle, more information becomes available (see Fig. 3.5) and it would appear that it is theoretically possible to always be able to obtain all the required values, r, m_r, m_i . This occurs for $r \geq 0.2\mu\text{m}$, although for particles with larger refractive indices ($m_r \geq 2.1$), the size of the particle need not be so big ($r \geq 0.1\mu\text{m}$). Physically, the phase function becomes more rapidly changing with larger size and refractive index, giving more complicated shapes. Whether a retrieval method will be able to do as well as is predicted here is not so clear.

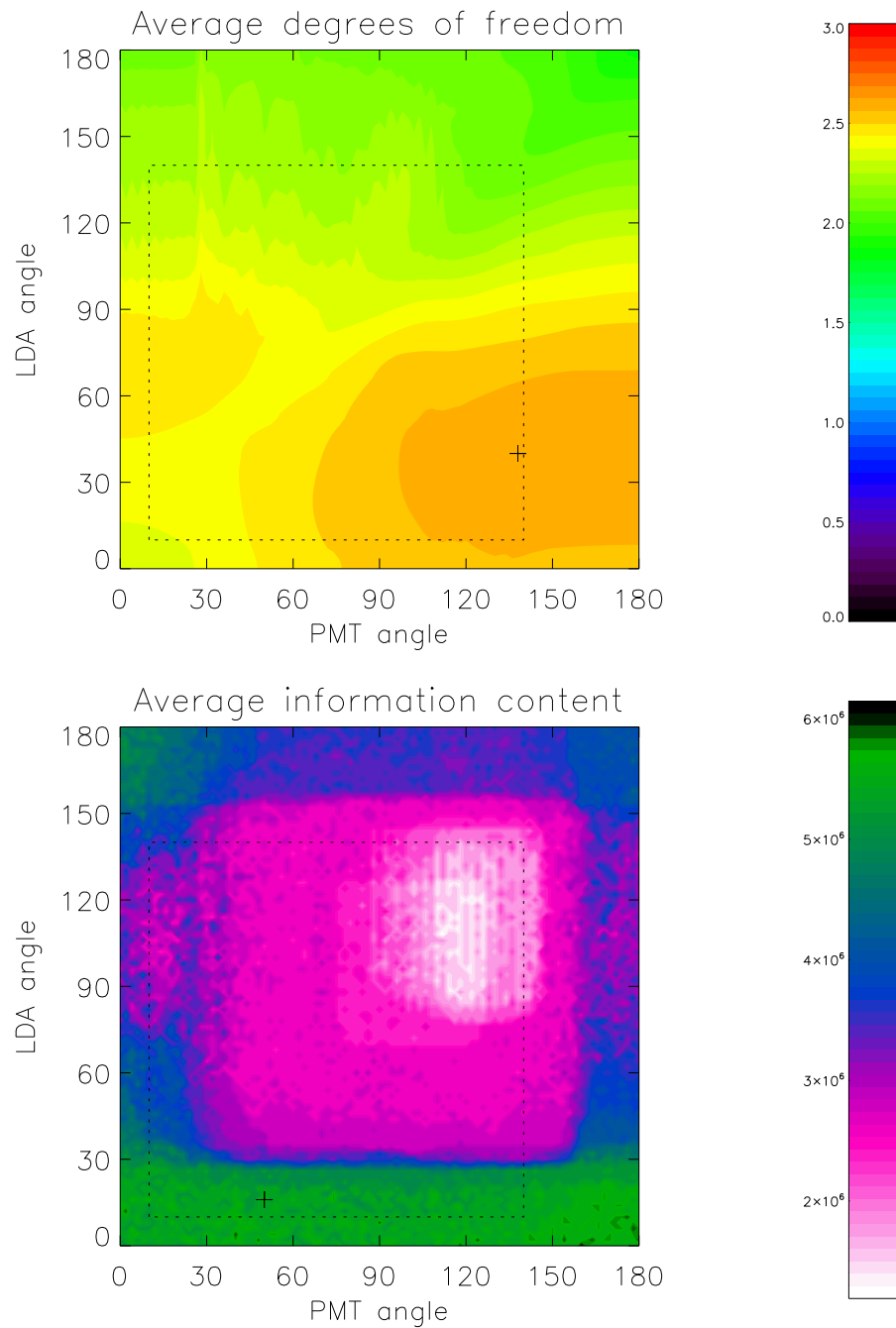


Figure 3.3: Optimisation calculations. The average degrees of freedom (d_s) and information content (H) for the range of r and m given in Table 3.1 are calculated and plotted as a function of detector angles. The dotted line shows the area in which the detectors are allowed to be placed. The cross shows the point with maximum average value of d_s . $\xi = 0^\circ$, detector distance = 30mm.

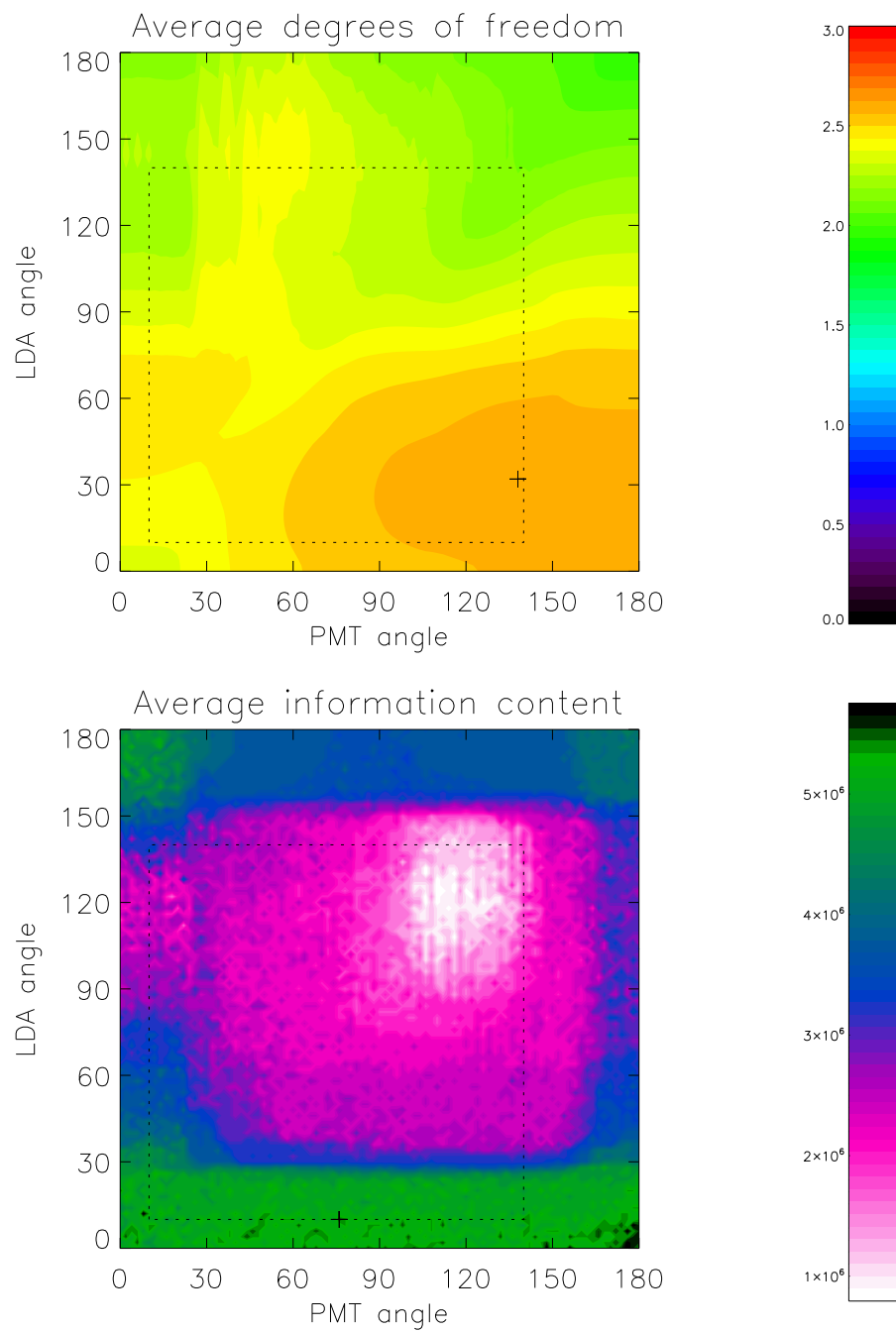


Figure 3.4: Average d_s and H for the range of r and m given in Table 3.1 plotted as a function of detector angles. $\xi = 90^\circ$, detector distance = 30mm.

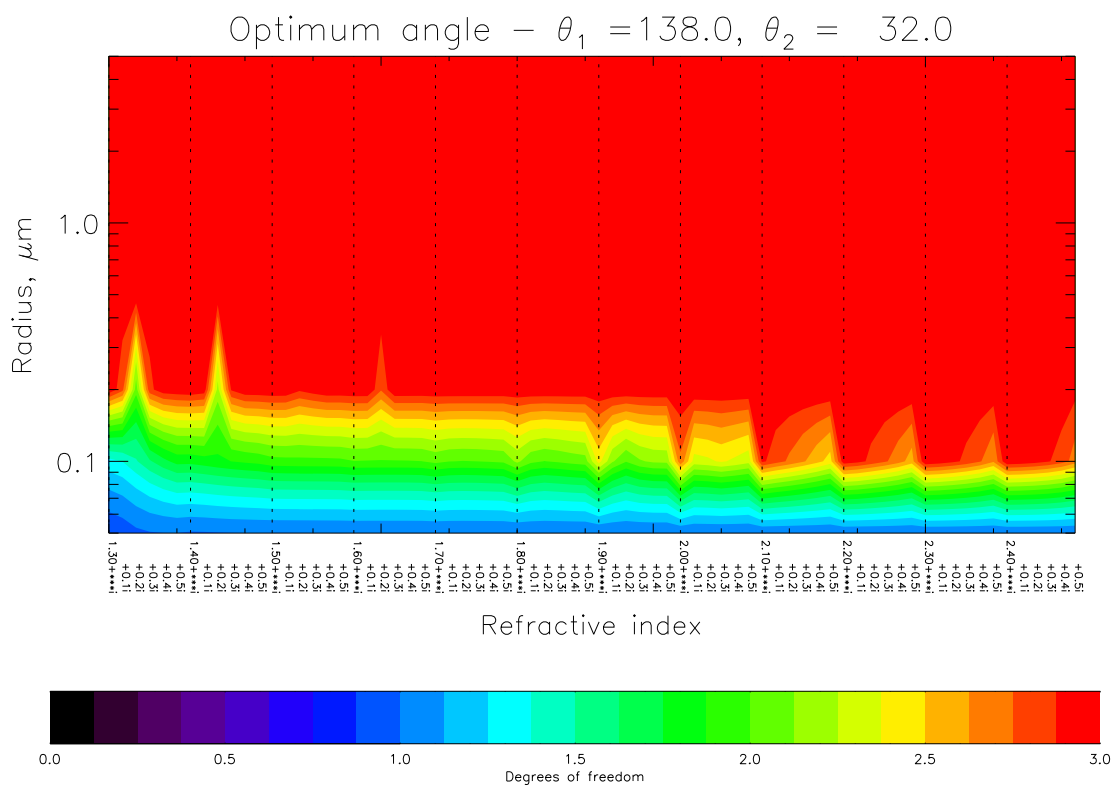


Figure 3.5: For the optimum angles of $\theta_{\text{PMT}} = 138^\circ$ and $\theta_{\text{LDA}} = 34^\circ$, the degrees of freedom for all calculated radii and refractive indices are shown. This plot is for $\xi = 90^\circ$, distance = 30mm, as in Fig. 3.4

Chapter 4

Time line and Future Work

4.1 Work to date

Michaelmas Term 2007

- Reacquainting with IDL and FORTRAN.
- Background reading on Mie theory and inverse methods.
- Familiarised myself with the EODG Mie routines [*Light scattering routines* 2008].
- Studied the AOPP 4th Undergraduate course and completed problem sets (this continued for the three terms).
- Began forward model for SPARCLE instrument.

Hilary Term 2008

- Extended the SPARCLE forward model, and wrote code to optimise the position of its detectors.
- Background reading in the T-matrix method.
- Attended workshop on non-spherical scattering in Bremen.
- Obtained Michael Mishchenko's FORTRAN T-matrix code [*T-Matrix Codes for Computing Electromagnetic Scattering by Nonspherical and Aggregated Particles* 2008] and modified, converting the main routines into callable procedures.

Trinity Term 2008

- Learnt to programme in C.
- Created IDL dynamically loadable modules (DLMs) to call Michael Mishchenko's T-matrix code in IDL.
- Attended APPRAISE science meeting at Leeds University and presented poster: *Optimising the positioning of sensors in the SPARCLE instrument.*

4.2 Plan for future work

4.2.1 Further work on SPARCLE

Now that parts have been ordered for the instrument, the exact geometries and specifications of components mean that the forward model for SPARCLE will have to be updated. Due to the modular form of the routines, this should not be too taxing. This means the optimisation calculations will also have to be recalculated to find the best angles for the specific devices purchased. The major difference between the originally proposed device and the ordered parts is that the linear diode array has been replaced with a 2D array. The forward model will also probably need to be optimised as it currently is written fairly inefficiently.

Hopefully, the optimisation method used will be able to form part of a paper on the design of SPARCLE.

4.2.2 Research into additional scattering methods

There are several more non-spherical scattering methods I would be interested in looking into. It will be useful to have a comparison of the speed and accuracy of different methods looking at identical scattering objects (although this will not be possible for certain schemes). The main additional methods I am interested in are:

Scattering of spheroids in a spheroidal coordinate system: This technique, by Asano & Yamamoto [1975], would presumably be much faster for spheroids than T-matrix theory. It is equivalent to Mie theory, in a different coordinate system although, as with other non-spherical methods, it is no longer possible to assume that the incoming light will approach along a preferred axis for convenience, and so calculations would surely be more tricky.

Null-field method with discrete sources: Doicu et al.'s [2006] solution to the T-matrix problem can be used to calculate much larger particles (and with more extreme geometries) than the standard solution given in §2.2. This is because the method uses multipole vector spherical expansions to represent the fields involved, with the expansions of the field from each pole matched in the overlapping regions. The use of multiple origins is helpful because highly a-spherical shapes can use expansions from close-by poles to reduce the number of expansion coefficients required to accurately portray the internal field in a particle. The downside of this method is the increased computer time required for calculations, but there is the significant advantage that the T-matrix of almost any shape can be calculated given enough time [Wriedt 2002].

Fractal aggregate methods: Mentioned in §2.3, I would be very interested in implementing some fractal methods of light scattering. If the results were adequately accurate, this would be a preferable method, as calculations should be rapid. For black carbon, it may be possible to model the evolution of the particles from creation to final form (as mentioned in §1.2.2) as an evolution in fractal dimension.

Compound scattering by spheres: Similar to the above T-matrix method, this technique uses multiple origins to expand VSWFs. In this case, each origin is the positioning of a new sphere, so that an aggregate of spheres is built up. I would be interested in using the code of Xu & Gustafson [n.d.] which can cope with spheres of different sizes at each local origin. Also, comparison with the cut down fractal method mentioned in the previous bullet point would be helpful to see if the full treatment is necessary.

Coated spheres: Essentially an extension of Mie theory to concentric scatterers, the idea of coated spheres can be applied sensibly to non-spherical atmospheric aerosol which have become coated for example, sand which has become surrounded by water, which will form an outer spherical shell. Well tested code by Bohren & Huffman [1983] is available, and could be compared against full non-spherical treatment by Doicu et al. [2006] who have written a solution for this case (which will take a far longer time to compute).

4.2.3 Analytical differentials and the T-matrix method

As far as I can tell, there are no published papers on the analytical differentiation of scattering characteristics obtained via the T-matrix method. These should be possible, since the calculations are analytic (with the exception of surface integrals in most cases). This would be a useful addition to

the differentials of spherical particle characteristics presented by this group [Grainger et al. 2004] and if rigorously derived, should be publishable although it could well take a significant amount of time.

4.2.4 Rapidity of calculations

While it is admirable to be able to calculate the scattering characteristics of horribly shaped particles to a good degree of accuracy, in the atmospheric sciences, the knowledge of scattered light is only a stepping stone in a larger calculation and as such, the use of non-spherical scattering code in forward models for atmospheric retrievals is prohibitively slow.

As a result, I intend to spend a large amount of time investigating practical methods for inserting non-spherical scattering code into forward models. This basically means obtaining efficient algorithms for presenting the data in a timely manner. Two ideas immediately come to mind:

Look up tables: These would allow computationally expensive calculations to be inserted into code using interpolation to obtain values in between table points.

Approximations: It may be possible to find approximations to full solutions similar to the Henyey-Greenstein function used as a replacement to the full Mie solution [Henyey & Greenstein 1941]. If errors between similar functions and a full solution were of the order 10% then this would be an acceptable alternative.

Bibliography

- Abel, S. J., Haywood, J. M., Highwood, E. J., Li, J. & Buseck, P. R. [2003], ‘Evolution of biomass burning aerosol properties from an agricultural fire in southern Africa’, *Geophysical Research Letters* **30**(15), 1783.
- Albrecht, B. A. [1989], ‘Aerosols, cloud microphysics, and fractional cloudiness’, *Science* **245**, 1227–1230.
- Arfken, G. B. & Weber, H. J. [1995], *Mathematical Methods for Physicists*, 4th edn, Academic Press. Very good section on the associated Legendre polynomials (pp722–735). There is also a comprehensive guide to Bessel functions (pp627–692).
- Asano, S. & Yamamoto, G. [1975], ‘Light scattering by a spheroidal particle’, *Applied Optics* **14**(1), 29–49.
- Baker, B. B. & Copson, F. T. [1987], *The mathematical theory of Huygen’s principle*, 3rd edn, AMS Chelsea Publishing.
- Bhutiyani, M. R., Kale, V. S. & Pawar, N. J. [2007], ‘Long-term trends in maximum, minimum and mean annual air temperatures across the Northwestern Himalaya during the twentieth century’, *Climatic Change* **85**(1), 159–177.
- Bluth, G. J. S., Doiron, S. D., Schnetzler, C. C., Krueger, A. J. & Walter, L. S. [1992], ‘Global tracking of the SO₂ clouds from the June, 1991 Mount Pinatubo eruptions’, *Geophysical Research Letters* **19**(2), 151–154.
- Boas, M. L. [1983], *Mathematical methods in the physical sciences*, 2nd edn, John Wiley & Sons. Chapter 12: The associated Legendre functions are discussed pp 504–506. Bessel functions are pp 509–526.
- Bohren, C. F. & Huffman, Donald, R. [1983], *Absorption and Scattering of Light by Small Particles*, Wiley-VCH. Chapter 4 has a very well explained derivation of Mie theory.

- Colbeck, I., ed. [1998], *Physical and Chemical Properties of Aerosols*, 1st edn, Blackie Academic & Professional.
- Doicu, A., Wriedt, T. & Eremin, Y. A. [2006], *Light Scattering by Systems of Particles: Null-Field Method with Discrete Sources - Theory and Programs*, Vol. 124 of *Optical Sciences*, Springer-Verlag, Berlin Heidelberg.
- Grainger, R. G., Lucas, J., Thomas, G. E. & Ewen, G. B. [2004], ‘Calculation of Mie derivatives’, *Applied Optics* **43**(28), 5386–5393.
- Haywood, J. M. & Shine, K. P. [1995], ‘The effect of anthropogenic sulfate and soot aerosol on the clear sky planetary radiation budget’, *Geophysical Research Letters* **22**(5), 603–606.
- Hecht, E. [2002], *Optics*, 4th edn, Addison Wesley, San Francisco. Gives brief outline of Ewald-Oseen Extinction Theorem and the points to a more thorough text.
- Heney, L. G. & Greenstein, J. L. [1941], ‘Diffuse radiation in the galaxy’, *The Astrophysical Journal* **93**, 70–83.
- Light scattering routines* [2008]. <http://www-atm.physics.ox.ac.uk/code/mie/> accessed on 27/8/2008. Details are in Grainger et al. [2004].
- Martins, J. V., Artaxo, P., Liousse, C., Reid, J. S., Hobbs, P. V. & Kaufman, Y. J. [1998], ‘Effects of black carbon content, particle size, and mixing on light absorption by aerosols from biomass burning in Brazil’, *Journal of Geophysical Research* **103**(D24), 32,041–32,050.
- Mergenthaler, J. L., Kumer, J. B. & Roche, A. E. [1995], ‘CLAES observations of Mt. Pinatubo stratospheric aerosol’, *Geophysical Research Letters* **22**(24), 3497–3500.
- Mie, G. [1908], ‘Beiträge zur Optik trüber Medien, speziell kolloidaler Metallösungen.’, *Annalen der Physik* **25**(3), 377–445. A translation can be found at <http://diogenes.iwt.uni-bremen.de/vt/laser/papers/RAE-LT1873-1976-Mie-1908-translation.pdf>.
- Mishchenko, M. I. [1990], ‘Extinction of light by randomly-oriented non-spherical grains’, *Astrophysics and Space Science* **164**(1), 1–13.
- Mishchenko, M. I. [2000], ‘Calculation of the amplitude matrix for a nonspherical particle in a fixed orientation’, *Applied Optics* **39**(6), 1026–1031.

- Mishchenko, M. I., Hovenier, J. W. & Travis, L. D., eds [2000], *Light Scattering by Nonspherical Particles: Theory, Measurements, and Applications*, Academic Press.
- Mishchenko, M. I. & Travis, L. D. [1998], ‘Capabilities and limitations of a current FORTRAN implementation of the T-matrix method for randomly oriented, rotationally symmetric scatterers’, *Journal of Quantitative Spectroscopy & Radiative Transfer* **60**(3), 309–324.
- Mishchenko, M. I., Travis, L. D. & Mackowski, D. W. [1996], ‘T-matrix computations of light scattering by nonspherical particles: A review’, *Journal of Quantitative Spectroscopy & Radiative Transfer* **55**(5), 535–575.
- Mishchenko, M. I., Videen, G., Khlebtsov, N. G., Wriedt, T. & Zakharova, N. T. [2008], ‘Comprehensive T-matrix reference database: A 2006–07 update’, *Journal of Quantitative Spectroscopy & Radiative Transfer* **109**(8), 1447–1460. Special edition for X Conference on Electromagnetic and Light Scattering by Non-Spherical Particles.
- Morse, P. M. & Feshbach, H. [1953], *Methods of Theoretical Physics, Pt. II*, McGraw-Hill, New York.
- NASA Earth Observatory News Room [2000]. http://earthobservatory.nasa.gov/Newsroom/NewImages/images.php3?img_id=1520, accessed on 3/7/2008.
- Peters, D. [2007], Calculation of SPARCLE’s signal to noise ratio, Technical report, Atmospheric, Oceanic and Planetary Physics, University of Oxford.
- Pincus, R. & Baker, M. B. [1994], ‘Effect of precipitation on the albedo susceptibility of clouds in the marine boundary layer’, *Nature* **372**(6503), 250–252.
- Reid, J. S. & Hobbs, P. V. [1998], ‘Physical and optical properties of young smoke from individual biomass fires in Brazil’, *Journal of Geophysical Research* **103**(D24), 32,013–32,030.
- Remer, L. A., Kaufman, Y. J., Holben, B. N., Thompson, A. M. & McNamara, D. [1998], ‘Biomass burning aerosol size distribution and modeled optical properties’, *Journal of Geophysical Research* **103**(D24), 31,879–31,891.
- Rodgers, C. D. [2000], *Inverse methods for atmospheric sounding: Theory and practice*, Vol. 2 of *Atmospheric, Oceanic and Planetary Physics*, World Scientific Publishing Co.
- Seinfeld, J. H. & Pandis, S. N. [1998], *Atmospheric Chemistry and Physics, From Air Pollution to Climate Change*, 1st edn, John Wiley & Sons, Inc.

- Solomon, S., Qin, D., Manning, M., Alley, R., Berntsen, T., Bindoff, N., Chen, Z., Chidthaisong, A., Gregory, J., Hegerl, G., Heimann, M., Hewitson, B., Hoskins, B., Joos, F., Jouzel, J., Kattsov, V., Lohmann, U., Matsuno, T., Molina, M., Nicholls, N., Overpeck, J., Raga, G., Ramaswamy, V., Ren, J., Rusticucci, M., Somerville, R., Stocker, T., Whetton, P., Wood, R. & Wratt, D. [2007], *Climate Change 2007: The Physical Science Basis*, Cambridge University Press. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change.
- Sorensen, C. M. [2001], 'Light scattering by fractal aggregates: A review', *Aerosol Science and Technology* **35**(2), 648–687.
- T-Matrix Codes for Computing Electromagnetic Scattering by Nonspherical and Aggregated Particles* [2008]. http://www.giss.nasa.gov/~crmim/t_matrix.html, accessed on 14/5/2008.
- Tsang, L., Kong, J. A. & Shin, R. T. [1985], *Theory of Microwave Remote Sensing*, Wiley Series in Remote Sensing, John Wiley & Sons, Inc. Section 5 in Chapter 3.
- Twomey, S. [1977], *Atmospheric Aerosols*, Vol. 7 of *Developments in Atmospheric Science*, Elsevier Scientific Publishing Company.
- van de Hulst, H. C. [1957], *Light Scattering by Small Particles*, John Wiley & Sons.
- Volten, H., Muñoz, O., de Haan, J. F., Vassen, W. & Hovenier, J. W. [2001], 'Scattering matrices of mineral aerosol particles at 441.6 nm and 632.8 nm', *Journal of Geophysical Research* **106**(D15), 17,375–17,401.
- Washington, R., Todd, M. C., Lizcano, G., Tegen, I., Flamant, C., Koren, I., Ginoux, P., Engelstaedter, S., Bristow, C. S., Zender, C. S., Goudie, A. S., Warren, A. & Prospero, J. M. [2006], 'Links between topography, wind, deflation, lakes and dust: The case of the Bodélé Depression, Chad', *Geophysical Research Letters* **33**, L09401.
- Waterman, P. C. [1965], Matrix formulation of electromagnetic scattering, in 'Proc. IEEE', Vol. 53, The Institute of Electrical and Electronics Engineers, Inc., pp. 805–812.
- Waterman, P. C. [1969], 'New formulation of acoustic scattering', *The Journal of the Acoustical Society of America* **45**(6), 1417–1429.
- Waterman, P. C. [1971], 'Symmetry, unitarity, and geometry in electromagnetic scattering', *Physical Review D* **3**(4), 825–839.

- Waterman, P. C. [1976], ‘Matrix theory of elastic wave scattering’, *Journal of the Acoustical Society of America* **60**(3), 567–580.
- Waterman, P. C. [1979], ‘Matrix methods in potential theory and electromagnetic scattering’, *Journal of Applied Physics* **50**(7), 4550–4566. The first derivation of T-matrix theory that did not rely on Green’s Dyadic and fictitious fields, leading to a much clearer explanation.
- Whitby, K. T. & Cantrell, B. K. [1976], Atmospheric aerosols: Characteristics and measurement, in ‘International Conference on Environmental Sensing and Assessment (ICESA)’, number 29–1, Institute of Electrical and Electronic Engineers (IEEE), ICESA, Las Vegas, NV, p. 6.
- Wiscombe, W. J. & Grams, G. W. [1976], ‘The backscattered fraction in two-stream approximations’, *Journal of the Atmospheric Sciences* **33**(12), 2440–2451.
- Wriedt, T. [2002], ‘Using the T-matrix method for light scattering computations by non-axisymmetric particles: Superellipsoids and realistically shaped particles’, *Particle & Particle Systems Characterization* **19**(4), 256–268.
- Xu, Y.-l. & Gustafson, B. A. S. [n.d.], An analytical solution to electromagnetic multisphere-scattering — the scatterign formulation used in codes gmm01f.f and gmm01s.f, Technical report, Department of Astronomy, University of Florida, Gainesville, P.O. Box 112055, FL 32611-2055.
- Yia, L., Koganb, Y. L. & Mechemc, D. B. [2008], ‘An idealized modeling study of the effect of continental air mass aerosol parameters on marine stratocumulus’, *Atmospheric Research* **88**(2), 157–167.

Appendix A

The expansion of vector spherical wave functions in spherical polar coordinates

The general VSWFs used in spherical scattering can be written in spherical coordinates as:

$$\mathbf{M}_{omn} = \frac{m}{\sin \theta} \cos m\phi P_n^m z_n \hat{\mathbf{e}}_\theta - \sin m\phi \frac{dP_n^m}{d\theta} z_n \hat{\mathbf{e}}_\phi, \quad (\text{A.1a})$$

$$\mathbf{M}_{emn} = \frac{-m}{\sin \theta} \sin m\phi P_n^m z_n \hat{\mathbf{e}}_\theta - \cos m\phi \frac{dP_n^m}{d\theta} z_n \hat{\mathbf{e}}_\phi, \quad (\text{A.1b})$$

$$\mathbf{N}_{omn} = \frac{z_n}{\rho} \sin m\phi n(n+1) P_n^m \hat{\mathbf{e}}_r + \sin m\phi \frac{dP_n^m}{d\theta} \frac{[\rho z_n]'}{\rho} \hat{\mathbf{e}}_\theta + m \cos m\phi \frac{P_n^m}{\sin \theta} \frac{[\rho z_n]'}{\rho} \hat{\mathbf{e}}_\phi, \quad (\text{A.1c})$$

$$\mathbf{N}_{emn} = \frac{z_n}{\rho} \cos m\phi n(n+1) P_n^m \hat{\mathbf{e}}_r + \cos m\phi \frac{dP_n^m}{d\theta} \frac{[\rho z_n]'}{\rho} \hat{\mathbf{e}}_\theta - m \sin m\phi \frac{P_n^m}{\sin \theta} \frac{[\rho z_n]'}{\rho} \hat{\mathbf{e}}_\phi. \quad (\text{A.1d})$$

The associated Legendre functions are $P_n^m(\cos \theta)$ and $z_n(\rho)$ are Bessel functions. Primes denote differentials with respect to the argument of z_n , ρ . Once we have simplified to the spherical case, with plane parallel incident light, we write (A.1) in terms of π_n and τ_n

$$\pi_n = \frac{P_n^1}{\sin \theta} \quad \text{and} \quad \tau_n = \frac{dP_n^1}{d\theta} :$$

$$\mathbf{M}_{o1n} = \cos \phi \pi_n z_n \hat{\mathbf{e}}_\theta - \sin \phi \tau_n z_n \hat{\mathbf{e}}_\phi, \quad (\text{A.2a})$$

$$\mathbf{M}_{e1n} = -\sin \phi \pi_n z_n \hat{\mathbf{e}}_\theta - \cos \phi \tau_n z_n \hat{\mathbf{e}}_\phi, \quad (\text{A.2b})$$

$$\mathbf{N}_{o1n} = n(n+1) \sin \phi \sin \theta \pi_n \frac{z_n}{\rho} \hat{\mathbf{e}}_r + \sin \phi \tau_n \frac{[\rho z_n]'}{\rho} \hat{\mathbf{e}}_\theta + \cos \phi \pi_n \frac{[\rho z_n]'}{\rho} \hat{\mathbf{e}}_\phi, \quad (\text{A.2c})$$

$$\mathbf{N}_{e1n} = n(n+1) \cos \phi \sin \theta \pi_n \frac{z_n}{\rho} \hat{\mathbf{e}}_r + \cos \phi \tau_n \frac{[\rho z_n]'}{\rho} \hat{\mathbf{e}}_\theta - \sin \phi \pi_n \frac{[\rho z_n]'}{\rho} \hat{\mathbf{e}}_\phi. \quad (\text{A.2d})$$

These equations are taken from Bohren & Huffman [1983]. One should be aware that these are **not** the VSWF used for T-matrix calculations in §2.2.

Appendix B

The orthogonality of vector spherical wave functions

Here, the integral

$$I = \iint dS \hat{\mathbf{n}} \cdot \left[(\nabla \times \mathbf{M}_{mn}^{(1)}) \times \mathbf{M}_{-mn}^{(3)} + \mathbf{M}_{mn}^{(1)} \times (\nabla \times \mathbf{M}_{-mn}^{(3)}) \right] \quad (\text{B.1})$$

is calculated. Very similar calculations apply to \mathbf{N} and the same result is obtained, but the mathematics is messier.

The integral can be carried out over any closed surface with the same result applying. For simplicity, a sphere of radius r is chosen so that, from (2.50), $dS \hat{\mathbf{n}} = r^2 \sin \theta d\theta d\phi \hat{\mathbf{e}}_r$. So (B.1) becomes:

$$I = \int_0^{2\pi} d\phi \int_0^\pi \sin \theta d\theta r^2 \hat{\mathbf{e}}_r \cdot \left[(\nabla \times \mathbf{M}_{mn}^{(1)}) \times \mathbf{M}_{-mn}^{(3)} + \mathbf{M}_{mn}^{(1)} \times (\nabla \times \mathbf{M}_{-mn}^{(3)}) \right]. \quad (\text{B.2})$$

In polar spherical coordinates, curl \mathbf{M} is written as

$$\begin{aligned} \nabla \times \mathbf{M} = & \frac{1}{r \sin \theta} \left[\frac{\partial}{\partial \theta} (M_\phi \sin \theta) - \frac{\partial M_\phi}{\partial \phi} \right] \hat{\mathbf{e}}_r \\ & + \frac{1}{r} \left[\frac{1}{\sin \theta} \frac{\partial M_r}{\partial \phi} - \frac{\partial}{\partial r} (r M_\phi) \right] \hat{\mathbf{e}}_\theta + \frac{1}{r} \left[\frac{\partial}{\partial r} (r M_\theta) - \frac{\partial M_r}{\partial \theta} \right] \hat{\mathbf{e}}_\phi, \quad (\text{B.3}) \end{aligned}$$

and from (2.26a) we obtain the components of \mathbf{M} :

$$\begin{aligned} M_r &= 0, \\ M_\theta &= \gamma_{nm} z_n(kr) e^{im\phi} \frac{im}{\sin\theta} P_n^m(\cos\theta), \\ M_\phi &= \gamma_{nm} z_n(kr) e^{im\phi} \frac{dP_n^m(\cos\theta)}{d\theta}. \end{aligned}$$

The only trick employed to obtain the derivatives is to say that

$$\frac{\partial}{\partial\theta}(M_\phi \sin\theta) = -\gamma_{nm} z_n e^{im\phi} \frac{d}{d\theta} \left[\sin\theta \frac{dP_n^m}{d\theta} \right] = \gamma_{nm} z_n e^{im\phi} \left[n(n+1) - \frac{m^2}{\sin^2\theta} \right] P_n^m \sin\theta, \quad (\text{B.4})$$

where we have used (2.5b). Plugging these components into (B.3) we obtain:

$$\nabla \times \mathbf{M}_{mn} = \frac{\gamma_{nm} e^{im\phi}}{r} \left\{ z_n(kr) n(n+1) P_n^m \hat{\mathbf{e}}_r + \frac{d}{dr} [r z_n(kr)] \left[\frac{dP_n^m}{d\theta} \hat{\mathbf{e}}_\theta + \frac{im}{\sin\theta} P_n^m \hat{\mathbf{e}}_\phi \right] \right\}. \quad (\text{B.5})$$

Next we consider $(\nabla \times \mathbf{M}_{mn}^{(1)}) \times \mathbf{M}_{-mn}^{(3)}$. Recall that $\hat{\mathbf{e}}_r \times \hat{\mathbf{e}}_\theta = \hat{\mathbf{e}}_\phi$ and $\hat{\mathbf{e}}_\theta \times \hat{\mathbf{e}}_\phi = \hat{\mathbf{e}}_r$. We also use (2.27) from §2.2 and (12.81a) from Arfken & Weber [1995, p724] to give us:

$$\gamma_{nm} \gamma_{n-m} = \gamma_{nm}^2 \frac{(n+m)!}{(n-m)!}, \quad P_n^{-m} = (-1)^m \frac{(n-m)!}{(n+m)!} P_n^m,$$

and we can then say

$$\begin{aligned} (\nabla \times \mathbf{M}_{mn}^{(1)}) \times \mathbf{M}_{-mn}^{(3)} &= \frac{-\gamma_{nm} \gamma_{n-m}}{r} \left\{ j_n n(n+1) P_n^m \hat{\mathbf{e}}_r + \frac{d}{dr} [r j_n(kr)] \left[\frac{dP_n^m}{d\theta} \hat{\mathbf{e}}_\theta + \frac{im}{\sin\theta} P_n^m \hat{\mathbf{e}}_\phi \right] \right\} \\ &\quad \times \left[h_n^{(1)} \left(\frac{im}{\sin\theta} P_n^{-m} \hat{\mathbf{e}}_\theta + \frac{dP_n^{-m}}{d\theta} \hat{\mathbf{e}}_\phi \right) \right], \end{aligned} \quad (\text{B.6})$$

$$\begin{aligned} &= \gamma_{nm}^2 \frac{(-1)^m}{r} \left\{ \left[\left(\frac{m P_n^m}{\sin\theta} \right)^2 + \left(\frac{dP_n^m}{d\theta} \right)^2 \right] \frac{d}{dr} [r j_n] h_n^{(1)} \hat{\mathbf{e}}_r \right. \\ &\quad \left. + j_n h_n^{(1)} n(n+1) P_n^m \left(-\frac{dP_n^m}{d\theta} \hat{\mathbf{e}}_\theta + \frac{im}{\sin\theta} P_n^m \hat{\mathbf{e}}_\phi \right) \right\}. \end{aligned} \quad (\text{B.7})$$

The expression $\mathbf{M}_{mn}^{(1)} \times (\nabla \times \mathbf{M}_{-mn}^{(3)})$ is easily found by multiplying (B.7) by -1 (since $\mathbf{a} \times \mathbf{b} = -\mathbf{b} \times \mathbf{a}$)

and switching j_n with $h_n^{(1)}$. Thus we find that

$$(\nabla \times \mathbf{M}_{mn}^{(1)}) \times \mathbf{M}_{-mn}^{(3)} + \mathbf{M}_{mn}^{(1)} \times (\nabla \times \mathbf{M}_{-mn}^{(3)}) = \frac{\gamma_{nm}^2 (-1)^m}{r} \left[\left(\frac{m P_n^m}{\sin \theta} \right)^2 + \left(\frac{dP_n^m}{d\theta} \right)^2 \right] \left[h_n^{(1)} \frac{d}{dr} [r j_n] - j_n \frac{d}{dr} [r h_n^{(1)}] \right] \hat{\mathbf{e}}_r, \quad (\text{B.8})$$

and we can write

$$I = r \gamma_{nm}^2 (-1)^m \left[h_n^{(1)} \frac{d}{dr} [r j_n] - j_n \frac{d}{dr} [r h_n^{(1)}] \right] \cdot \int_0^{2\pi} d\phi \int_0^\pi \sin \theta d\theta \left[\left(\frac{m P_n^m}{\sin \theta} \right)^2 + \left(\frac{dP_n^m}{d\theta} \right)^2 \right]. \quad (\text{B.9})$$

The integral over ϕ is simple and the solution to the integral over θ is found as an example in Arfken & Weber [1995, p734]. The only thing left to wonder about is the value of the radial functions in the brackets. Since the calculation can be carried out at any radius, we use the $r \rightarrow \infty$, asymptotic values of the Bessel functions [Arfken & Weber 1995, Bohren & Huffman 1983] which are:

$$h_n^{(1)}(kr) \sim \frac{(-i)^n e^{ikr}}{i kr}, \quad j_n(kr) \sim \frac{1}{kr} \sin(kr - \frac{n\pi}{2}), \quad (\text{B.10})$$

$$\text{which gives} \quad (\text{B.11})$$

$$\frac{d}{dr} [r h_n^{(1)}] \sim (-i)^n e^{ikr}, \quad \frac{d}{dr} [r j_n] \sim \cos(kr - \frac{n\pi}{2}). \quad (\text{B.12})$$

And so we discover that

$$I = r \gamma_{nm}^2 (-1)^m \frac{1}{i kr} 2\pi \cdot \frac{2n(n+1)}{2n+1} \cdot \frac{(n+m)!}{(n-m)!} = \gamma_{nm}^2 (-1)^m \frac{1}{ik} \frac{1}{\gamma_{nm}^2}, \quad (\text{B.13})$$

$$= \frac{(-1)^m}{i k}. \quad (\text{B.14})$$