OXFORD

POLARISATION

S. R. CLOUDE

POLARISATION: APPLICATIONS IN REMOTE SENSING

This page intentionally left blank

Polarisation Applications in Remote Sensing

S. R. CLOUDE



OXFORD UNIVERSITY PRESS

Great Clarendon Street, Oxford OX2 6DP

Oxford University Press is a department of the University of Oxford. It furthers the University's objective of excellence in research, scholarship, and education by publishing worldwide in

Oxford New York

Auckland Cape Town Dar es Salaam Hong Kong Karachi Kuala Lumpur Madrid Melbourne Mexico City Nairobi New Delhi Shanghai Taipei Toronto

With offices in

Argentina Austria Brazil Chile Czech Republic France Greece Guatemala Hungary Italy Japan Poland Portugal Singapore South Korea Switzerland Thailand Turkey Ukraine Vietnam

Oxford is a registered trade mark of Oxford University Press in the UK and in certain other countries

Published in the United States by Oxford University Press Inc., New York

© S. R. Cloude 2010

The moral rights of the author have been asserted Database right Oxford University Press (maker)

First Edition 2010

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, without the prior permission in writing of Oxford University Press, or as expressly permitted by law, or under terms agreed with the appropriate reprographics rights organization. Enquiries concerning reproduction outside the scope of the above should be sent to the Rights Department, Oxford University Press, at the address above

You must not circulate this book in any other binding or cover and you must impose the same condition on any acquirer

British Library Cataloguing in Publication Data

Data available

Library of Congress Cataloging in Publication Data

Cloude, Shane.
Polarisation : applications in remote sensing / S.R. Cloude.
p. cm.
ISBN 978-0-19-956973-1 (hardback)
1. Electromagnetic waves—Scattering.
2. Polarimetric remote sensing.
3. Interferometry.
I. Title.
QC665.S3C56
2009
539.2-dc22
2009026998

Typeset by Newgen Imaging Systems (P) Ltd., Chennai, India Printed in Great Britain on acid-free paper by CPI Antony Rowe, Chippenham, Wiltshire

ISBN: 978-0-19-956973-1 (Hbk.)

10 9 8 7 6 5 4 3 2 1

Preface

An alternative title considered for this book was *Which Way is Up? Questions and Answers in Polarisation Algebra.* On advice it was rejected in favour of a more conventional approach. Still, it is a good question. Which way is up? A question with a literal scientific interpretation—namely, how to define vertical in a free reference frame for electromagnetic waves, but also one with a colloquial interpretation about the best route to progress. At a technical level this book is concerned with the answer to the former, but hopefully will serve to promote in the reader some idea of the latter. It arises from over twenty years' personal experience of research in the topic, but also through the privilege of having met and collaborated with many of those who made fundamental contributions to the subject. Much of this original work remains, unfortunately, scattered in the research literature over different years and journals. This book, then, is an attempt to bring it all together in a didactic and coherent form suitable for a wider readership.

The book aims to combine—I believe for the first time—the topics of wave polarisation and radar interferometry, and to highlight important developments in their fusion: polarimetric interferometry. Here indeed we shall see that the whole is greater than the sum of the parts, and that by combining the two we open up new possibilities for remote sensing applications.

It is intended as a graduate level text suitable for a two-semester course for those working with radar remote sensing in whatever context, but is also aimed at working scientists and engineers in the broad church that is remote sensing. Hopefully it will also appeal to those working in optical physics—especially polarimetry and light scattering—and to mathematicians interested in aspects of polarisation algebra.

Before reviewing the structure of the book, certain spelling requires clarification. Polarisation or Polarization? The usual response is that British English uses 's', and American 'z'. However, in this text we reserve spelling with 's' for the property of a transverse wave, while we use 'z' for the effect of electromagnetic fields on matter. Hence waves remain polarised while matter is polarized. In this way we take advantage of both forms.

Chapter 1 first provides an introduction to the physical properties of polarised waves using the formal machinery of electromagnetic wave theory. The idea is to provide motivation and a foundation for many concepts used in later chapters. For example, the concepts of matrix decomposition, the use of the Pauli matrices in wave propagation and scattering and, most importantly of all, the idea of using unitary matrices to form a bridge between mathematical descriptions of polarisation in terms of complex and real numbers, are all introduced in this chapter. This is in addition to the more prosaic elements of polarisation theory, such as the polarisation ellipse, the Stokes vector, and the Poincaré sphere, all of which are covered. The chapter is organized around three main themes: how

to generate polarised waves and describe them in various coordinate systems, how to represent the propagation of such waves between two points A and B, and finally how to describe their interaction with particles via the process of scattering. The idea throughout is to develop the concept of the 'memory' imprinted on a wave of its original polarisation and how this may be lost through the complexities of propagation and scattering.

This idea of 'loss of memory' is developed further in Chapter 2, where stochastic effects are treated in more detail. We start by considering the coherency matrix of a wave and show how it leads to the wave dichotomy; namely, two different ways in which to model the loss of polarisation information to noise. This then opens up a new approach to describing the effects of noise, not just on a freely propagating wave but also on a scattering system as a whole via the concept of scattering entropy. Entropy is an important concept in this book and here we show how entropy from a generalized coherency matrix description can be formally linked to the classical Mueller/Stokes formulation. This leads, for example, to a formal test for isolating the set of physical Mueller matrices from the much wider set of 4×4 real matrices—something which is quite difficult to do from the Mueller calculus itself. We also show how the entropy concept can be applied to multiple dimensions, including general bistatic or forward scattering, so freeing it from the important but special case of backscatter widely used in radar.

Chapter 3 was in many ways one of the most difficult to write. Here we attempt to apply the ideas of entropy to electromagnetic models of surface and volume scattering (where polarization becomes important). What makes it difficult is the sheer scope of the problem. There are so many such models that they perhaps deserve a whole book to themselves. Instead we concentrate on a few simple models to convey the key ideas, and also link to developments in later chapters on decomposition theory and interferometry. Given that the main application of this book is to microwave scattering, we further concentrate on low-frequency models, whereby the wavelength is quite large compared to the size of the scattering feature, which has the further advantage that closed-form analytic formulae are available to calculate, for example, the scattering entropy. Having discussed this, we provide some treatment of high-frequency models and how they differ in polarisation properties from the low-frequency approach.

Chapter 4 deals with the important new topic of decomposition theorems. These now have widespread application in microwave remote sensing, and basically seek to isolate or separate various contributions in a mixture of scattering processes. The most important such idea is to separate surface from volume scattering. Microwaves have the ability to penetrate vegetation and other land cover (snow, ice, and so on) and thus generally incorporate a complicated mixture of processes in the scattered signal. Decomposition theorems are an attempt to separate these and hence improve interpretation and parameter retrieval in quantitative remote sensing applications. There are two basic classes of decomposition–coherent and incoherent–and within each class several authors have proposed different models. Here we provide a unified survey of all such methods and illustrate their various strengths and weaknesses by linking their physical structure to the ideas developed in earlier chapters.

One key conclusion we will see from the first four chapters is that entropy or 'loss of memory' about polarisation is often linked directly to the randomness of the scattering medium, and that the remote sensing 'observer' has little control over this. This is problematic for applications, for example, in vegetation remote sensing, where randomness in the volume leads to loss of polarisation information. A key idea for the second part of the book is therefore how to achieve some kind of entropy control in remote sensing of random media. One way to do this is to employ interferometry. Radar interferometry is a mature established topic, so in Chapter 5 we provide only a brief introduction for those not familiar with the key concepts. However, the chapter also contains one or two novel developments required in later chapters. In particular we develop a Fourier–Legendre series approach to a description of coherent volume scattering in interferometry. This then provides a bridge between the two halves of the book, and allows us to consider, in Chapter 6, the combination of polarisation diversity with interferometry.

The combination of polarisation diversity with radar interferometry has been a key development over the past decade. It was first made possible from an experimental point of view by late additions to the NASA Shuttle imaging radar mission SIR-C in 1994, and since then has evolved through a combination of theoretical studies and airborne radar experiments. In Chapter 6 we outline the basic theory of the topic, showing how to form interferograms in different polarisation channels before considering mathematically the idea of coherence optimization, whereby we seek the polarisation that maximizes the coherence (or minimizes the entropy). In this way we provide a link with earlier chapters by showing how polarimetric interferometry leads to a form of 'entropy control', even in random media applications.

In Chapter 7 we therefore revisit the ideas of surface and volume scattering first introduced in Chapter 4, but this time we investigate their properties in both interferometry and polarimetry. This is built around the idea of a coherence loci, a geometrical construct to bound the variation of interferometric coherence with polarisation, and closely related to the coherence region, the latter taking into account spread due to statistical estimation of coherence from data. Given the importance of surface/volume decompositions in microwave remote sensing, we treat in some detail the two-layer scattering problem of a volume layer on top of a surface and use it to review several model variations that are found in the literature.

In Chapter 8 we use these ideas to investigate the inverse problem: the estimation of model parameters from observed scattering data. We concentrate on the two-layer geometry and investigate four classes of problem. We start with the simplest: estimation of the lower bounding surface position, which is a basic extension of conventional interferometry and allows us, for example, to locate surface position beneath vegetation and hence remove a problem called vegetation bias in digital elevation models (DEMs). We then look at estimating the top of the layer, which corresponds in vegetation terms to finding forest height. This is an important parameter for estimating forest biomass, for example, and in assessing the amount of carbon stored in above-ground vegetation. We then look at the possibility of imaging a hidden layer using polarimetric interferometry. In this case we wish to filter out the scattering from a volume layer to image a surface beneath. The next logical step is to image the vertical variation of scattering through the layer itself, and this we treat as the topic of polarisation coherence tomography or PCT, which combines the Fourier–Legendre expansion of coherent volume scattering with decomposition theory in an interesting example of what can happen when two of the major themes of this book—polarisation and interferometry—are fused.

Finally, in Chapter 9 we turn attention to illustrative examples of these theoretical concepts. By far the most important current application area is in radar imaging or synthetic aperture radar (SAR), and so we begin by reviewing the basic concepts behind this technology, always highlighting those issues of particular importance to polarisation. We treat a hierarchy of such imaging systems, from SAR to POLSAR and POLINSAR, and then consider illustrative current applications in surface, volume, and combined surface and volume scattering.

We then present supportive material in three Appendices. In the first we provide a basic introduction to matrix algebra. This is used extensively in descriptions of polarised wave scattering, and is provided here to help those not familiar with the terminology and notation employed.

As mentioned earlier, one key idea in this book is the role played by unitary matrix transformations in linking (or mapping) different representations of polarisation algebra. For this reason, in Appendix 2 we provide a detailed mathematical treatment of the algebra behind such relationships, introducing concepts from Lie algebra, group theory, and matrix transformations to illustrate the fundamental relationships between complex and real representations of polarised wave scattering.

Finally, in Appendix 3 we provide a short treatment of stochastic signal theory as it relates to polarisation and interferometry. Here we treat aspects of speckle noise in coherent imaging, and show how estimation errors impact on estimation of scattered field parameters in remote sensing.

This book is the culmination of many years of study and research, and acknowledgement must be given to those many colleagues and students who provided the impetus and curiosity to study and develop these topics. Acknowl-edgements and thanks are extended to the European Microwave Scattering Laboratory (EMSL) at Ispra, Italy, for their permission to use data from their large anechoic chamber facility; to the German Aerospace Centre (DLR) in Oberpfaffenhofen, Germany, for provision of airborne radar data from their E-SAR system; and to Michael Mishchenko of NASA Goddard Space Center, USA, for provision of his latest numerical simulations of multiple scattering from particle clouds. Thanks also to the Japanese Space Agency (JAXA) for provision of the PALSAR satellite data used in Chapter 9. All these datasets play a vital role in illustrating the theory outlined in this book, and, I believe, help enormously in clarifying what would otherwise remain abstract concepts.

Key personal thanks go to five colleagues in particular. Firstly, to Professor Wolfgang Boerner of the University of Illinois, Chicago, USA. His early vision and boundless energy have inspired several generations of researchers in these topics, including my own early studies as a PhD student. Secondly, thanks to Professor Eric Potter of the University of Rennes, France. Our early collaboration on radar polarimetry, and particularly on decomposition theory, was inspiring, and has lead, I am pleased to say, to a lifelong friendship and collaboration. Thanks also to Drs Irena Hajnsek and Kostas Papathanassiou of the German Aerospace Centre, DLR. Their support and their contributions to the development of polarimetric interferometry have been key in the maturation of the subject. Finally, however, I would like to acknowledge the late Dr. Ernst Luneburg of DLR, Germany. His combination of scholarship and passion for the application of mathematics to remote sensing was the true inspiration for me to write this book, and I feel I can now finally answer his oft-posed question: *'Wo ist das Buch?'*

Shane Cloude January 2009

This page intentionally left blank

Contents

1	Polar	ised electromagnetic waves	1	
	1.1	The generation of polarised waves	3	
	1.2	The propagation of polarised waves	9	
	1.3	The geometry of polarised waves	34	
	1.4	The scattering of polarised waves	46	
	1.5	Geometry of the scattering matrix	60	
	1.6	The scattering vector formulation	67	
2	Depo	larisation and scattering entropy	71	
	2.1	The wave coherency matrix	72	
	2.2	The Mueller matrix	78	
	2.3	The scattering coherency matrix formulation	85	
	2.4	General theory of scattering entropy	91	
	2.5	Characterization of depolarising systems	110	
	2.6	Relating the Stokes/Mueller and coherency matrix		
		formulations	113	
3	Depo	larisation in surface and volume scattering	115	
	3.1	Introduction to surface scattering	116	
	3.2	Surface depolarisation	133	
	3.3	Introduction to volume scattering	142	
	3.4	Depolarisation in volume scattering	155	
	3.5	Simple physical models for volume scattering and	167	
		propagation	107	
4	Deco	mposition theorems	178	
	4.1	Coherent decomposition theorems	178	
	4.2	Incoherent decomposition theorems	189	
5	Introduction to radar interferometry		208	
	5.1	Radar interferometry	208	
	5.2	Sources of interferometric decorrelation	219	
6 Polarimetric interferometry				
	6.1	Vector formulation of radar interferometry	234	
	6.2	Coherence optimization	240	

7	The	coherence of surface and volume scattering	252
	7.1	Coherence loci for surface scattering	253
	7.2	Coherence loci for random volume scattering	255
	7.3	The coherence loci for a two-layer	
		scattering model	265
	7.4	Important special cases: RVOG, IWCM	
		and OVOG	270
8	Para	meter estimation using polarimetric	
	interf	ferometry	284
	8.1	Surface topography estimation	284
	8.2	Estimation of height $h_{\rm v}$	295
	8.3	Hidden surface/target imaging	315
	8.4	Structure estimation: extinction and Legendre	
		parameters	322
9	Appli	ications of polarimetry and interferometry	340
	9.1	Radar imaging	340
	9.2	Imaging interferometry: InSAR	345
	9.3	Polarimetric synthetic aperture radar (POLSAR)	347
	9.4	Polarimetric SAR interferometry (POLInSAR)	360
	9.5	Applications of polarimetry and interferometry	363
Appendix 1 Introduction to matrix algebra			
Appendix 2 Unitary and rotation groups			
Appendix 3 Coherent stochastic signal analysis			
Bibliography			435
Index			451

Polarised electromagnetic waves

The term 'wave polarisation' is relatively recent in the history of optics. It was first used by Étienne Malus (1775–1812) in 1809, although the 'orientability' of optical waves was certainly known by Isaac Newton (1643–1727) and Christiaan Huygens (1629–1695). They were concerned with a description of the strange phenomenon of double refraction in Iceland spar (calcite), first presented by Rasmus Bartholin (1625–1698) in 1670, and an explanation was set to challenge the best minds in optics for the ensuing 150 years. (For an introduction to the historical importance of polarisation in optics and its role in nature, see Collet, 1993; Iniesta, 2003; Konnen, 1985.)

It was, however, Thomas Young (1773–1829) who first suggested, in 1817, that polarisation may arise due to a transverse wave component of light-a controversial suggestion at the time, but an idea that was further developed and quantified by Augustin-Jean Fresnel (1788-1827) in 1821, with the development of the Fresnel equations for polarisation by surface reflection. This was followed in 1852 by the development of a mathematical theory of partially polarised waves by George Gabriel Stokes (1819–1903), based on the concept of a four-element Stokes vector (Stokes, 1852). However, it was only with the development of the electromagnetic wave theory of James Clerk Maxwell (1831–1879) in 1861 that light and indeed all electromagnetic waves were formally shown to be transverse and thus 'carry a memory of orientation' in propagating from source to observer (Jones, 1989). The reader should note, however, that Maxwell's theory caused some controversy at the time, and an interesting and readable account of the (sometimes turbulent) evolution of what we now call Maxwell's equations can be found in Hunt (1991).

In this book we concentrate on this orientation 'memory effect' and investigate ways in which it can be used for remote sensing. In a more general sense, this can be considered a subset of the wider, more formal topic of vector electromagnetic inverse problems (Boerner, 1981, 1992; Hopcraft, 1992).

In the post-Maxwell era there were four main developments of historical interest in the description of polarised waves. Firstly we mention the work of Henri Poincaré (Poincaré, 1892), who formalized many useful concepts in polarisation optics using a strongly geometrical approach. This was followed in 1941 by the first use of formal matrix algebra to describe the propagation of vector waves, by R. Clark Jones of the Polaroid Corporation and Harvard University. At about the same time, Hans Mueller, at the Massachussetts Institute of Technology, developed a matrix calculus for dealing with *partially* polarised



Fig. 1.1 A tripartite decomposition of active remote sensing systems

waves. In the radar community, early application of matrix algebra to scattering was carried out by Edward Kennaugh at Ohio State University. Finally, the concept of coherency matrices, first developed by Norbert Wiener in 1930, were first applied to polarisation algebra by Emil Wolf in 1954, and in 1960 Parrant and Roman formally linked polarisation algebra to the density matrix of statistical quantum mechanics. However, the coherency matrix formulation has much wider applicability to polarisation algebra than was originally foreseen, and in this book we explore this relationship in more detail and provide an updated treatment of these concepts.

Before treating these advanced topics, however, in this first chapter we use the machinery of electromagnetic wave theory to consider the basic mechanisms behind generation, propagation, and scattering of polarised electromagnetic waves. The formalism so developed will allow us to propagate a wave from the source to a scattering object and back again, so forming a basic template for the treatment of active remote sensing systems. Figure 1.1 shows a schematic representation of this tripartite decomposition of wave problems. We shall follow the logical progression of the diagram and begin with a description of the generation of polarised waves. We start with a general, coordinate free description based on the vector form of Maxwell's equations (Chen, 1985) before quickly focusing on three important coordinate systems, first classified in antenna theory in Ludwig (1973), and now widely used in analysis, engineering measurements and physical modelling. From these we can then define the concept of co- and crosspolarised fields. and ask the basic question as to whether the perfectly polarised source exists, even theoretically. (For the answer, see equation (1.16)and subsequent discussion.)

Having described how to generate polarised waves we then introduce vector wave propagation. This is a major topic in itself, and so in order to quickly bring forward the main ideas we require later in this book, we proceed by considering three specific examples. We start with the simplest—wave propagation in homogenous isotropic media—before examining two more exotic cases, where we will see how the concept of wave orthogonality can be formally defined and the family of polarisation types extended to include elliptical and circular polarisations.

Finally we consider the complex process of wave scattering, whereby secondary currents are induced in an object by the incident field and act as new sources of radiation to transfer information about the scatterer back to the observer. This process is characterized in the far field (that is, for large separations of source and object) for all polarisation states by a complex scattering amplitude matrix [S], the measurement and analysis of which forms a central theme of this book.

1.1 The generation of polarised waves

1.1.1 Maxwell's equations and vector plane waves

Electromagnetic waves are generated by accelerating charges (Jackson, 1999; Cloude, 1995a). The time and space variation of electric and magnetic fields are governed by a set of four partial differential equations, called Maxwell's equations, which can be written succinctly in the form shown on the left in equation (1.1) (Chen, 1985; Born and Wolf, 1989; Ishimaru 1991):

$$\nabla \times \underline{E} = -\frac{\partial \underline{B}}{\partial t}$$

$$\nabla \times \underline{H} = \underline{J} + \frac{\partial \underline{D}}{\partial t}$$

$$\overline{\underline{B}} = \mu_0 \underline{\underline{H}} \xrightarrow{\text{wave equation}} \nabla \times \nabla \times \underline{E} + \varepsilon_0 \mu_0 \frac{\partial^2 \underline{E}}{\partial t^2} = -\mu_0 \frac{\partial \underline{J}}{\partial t}$$

$$\nabla \cdot \underline{\underline{D}} = \rho$$

$$(1.1)$$

It is an interesting consequence of Maxwell's equations that even vacuum or free space is characterized by a pair of important constants: the permeability μ_0 and permittivity ε_0 , which have values derived from experiment, as shown in equation (1.2):

$$\mu_0 = 4\pi \times 10^{-7} H/m \quad \varepsilon_0 = 8.854 \times 10^{-12} F/m \tag{1.2}$$

Equation (1.1) then relates the radiated vector fields \underline{E} , \underline{B} , \underline{D} , and \underline{H} to source vector currents \underline{J} and scalar charge density ρ . The explicit differential equations relating these quantities can then be derived from equation (1.1) by treating the ∇ operator as a vector of partial derivative operations and using the following results from linear algebra:

$$\nabla = \begin{pmatrix} \frac{\partial}{\partial x}, & \frac{\partial}{\partial y}, & \frac{\partial}{\partial z} \end{pmatrix} \quad \underline{a} \times \underline{b} = \begin{vmatrix} \underline{i} & \underline{j} & \underline{k} \\ a_x & a_y & a_z \\ b_x & b_y & b_z \end{vmatrix}$$
(1.3)

The cross-product of ∇ with a vector is called the 'curl' (or sometimes 'rot' for rotor) operator, and the dot product the divergence or 'div'. In this book we shall be primarily concerned with the 'memory' these fields have for the vector nature of their source (that is, its orientation in space and structure in time), and how this may be used for remote sensing purposes.

The vector currents \underline{J} and scalar charges ρ are sources of the fields in equation (1.1). To demonstrate these as equivalent to a time derivative of current, we generate a vector wave equation by first forming a secondary vector product as $\nabla \times \nabla \times \underline{E}$ and then using the $\nabla \times \underline{H}$ Maxwell equation plus constitutive relations to eliminate \underline{B} . The result is shown on the right-hand side of equation (1.1). Note that on the left of this equation we have mixed second time and space derivatives of the electric field vector, while on the right we have the source of these fields, localized as the time derivative of vector currents. As current itself is caused by the time derivative of charge, it follows that radiation is caused by the second time derivative or charge acceleration. This acceleration is a vector quantity, the orientation of which is transferred into the radiated fields in the

form of propagating waves. While the form of these waves can be very general (see Cloude, 1995a), it is useful to start with a special type of solution: namely, vector plane waves.

For plane wave solutions we postulate electric field and driving current vectors of the form shown in equation (1.4). By adopting these simple plane wave solutions the space and time derivatives take on the simplified form shown on the right-hand side of this equation:

$$\frac{\underline{E} = \underline{E}_0 e^{i\left(\omega t - \underline{\beta} \cdot \underline{r}\right)}}{\underline{J} = \underline{J}_0 e^{i\omega t}} \right\} \Rightarrow \frac{\frac{\partial}{\partial t} \equiv i\omega}{\nabla \equiv -i\underline{\beta}}$$
(1.4)

Where $\omega = 2\pi f$ and $\beta = 2\pi/\lambda$, *f* is the frequency of the wave in Hertz, λ its wavelength in metres, and throughout this text we set $i = \sqrt{-1}$. Note that our notation, with a positive sign for the time derivative, is chosen by convention and leads to a complex refractive index for lossy material with a negative imaginary part (see Section 3.1.1.1). Be aware, however, that other notations exist in the literature, with some authors choosing \underline{E}^* for the plane wave, which changes the sign of the time derivative and leads to a complex refractive index with positive imaginary part (and which also impacts on the sense of circular polarisations, as we shall see). By direct substitution we then find that the vector wave equation in (1.1) has a solution when $|\underline{\beta}| = \beta = \sqrt{\omega^2 \varepsilon_0 \mu_0}$, and we obtain a vector Helmholtz equation of the form shown in equation (1.5), where we have now eliminated explicit time dependence.

$$\nabla \times \nabla \times \underline{E} - \beta^2 \underline{E} = i\omega \mu_o \underline{J} \tag{1.5}$$

The importance of such simple vector plane waves follows from the linearity of Maxwell's equations since, by superposition, the field at any location \underline{x} can then be obtained as a sum of contributions from all the source currents at locations \underline{y} . Hence we can express the solution of equation (1.5) as an integral or sum of the form shown in equation (1.6) (Chen, 1985).

$$\underline{\underline{E}}(\underline{x}) = i\omega\mu_0 \int\limits_{V} \overline{\overline{\underline{G}}}(\underline{x},\underline{y}) \cdot \underline{J}(\underline{y}) dV$$
(1.6)

The propagator of vectors from \underline{y} to \underline{x} is termed the dyadic Green's function, and by formally solving the vector Helmholtz equation for a Dirac delta source (a point source in space) it can be shown to have the following general form (Chen, 1985):

$$\overline{\overline{G}}(\underline{x},\underline{y}) = \left(\overline{\overline{I}} - \underline{r}\underline{r}\right)g + \frac{i}{kR}(\overline{\overline{I}} - 3\underline{r}\underline{r})g - \frac{1}{k^2R^2}(\overline{\overline{I}} - 3\underline{r}\underline{r})g$$
(1.7)

Here $\overline{\overline{I}}$ is the 3 × 3 unit dyad and has the form of a unit matrix, $R = |\underline{x} - \underline{y}|$, $\underline{r} = (\underline{x} - \underline{y})/|\underline{x} - \underline{y}|$, and the scalar Green's function g accounts for causality

and energy conservation as shown in equation (1.8):

$$g(\underline{x}, \underline{y}) = \frac{e^{-i\beta \left|\underline{x} - \underline{y}\right|}}{4\pi \left|\underline{x} - \underline{y}\right|}$$
(1.8)

As we move further away from the source currents, then $R \to \infty$ and the first term of $\overline{\overline{G}}$ dominates. Hence in the far-field, the dyadic Green's function simplifies by definition to the following form:

$$\overline{\overline{G}}_{\infty} = \left(\overline{\overline{I}} - \underline{r}\,\underline{r}\right)\frac{e^{-i\beta R}}{4\pi R} \tag{1.9}$$

The first part of this expression shows that only the components of \underline{J} transverse to the direction of propagation \underline{r} contribute to the radiated field. It follows from this that the radiated fields are transverse to the direction of propagation of the wave (called transverse electromagnetic or TEM waves). The electric field is defined from an integral sum over all currents, but the resultant must always lie in a plane perpendicular to \underline{r} . This is called the plane of polarisation and the resultant time locus of the electric field in this plane, the polarisation of the radiated wave.

To illustrate this, consider the fields radiated by elementary dipoles. In electromagnetic theory there are two types to consider: electric and magnetic (Jackson, 1999). For the electric dipole, current is localized at the origin, and an electric dipole moment p_0 generates an effective current distribution of the form shown in equation (1.10). Now evaluating the integral using the far field Green's dyadic (equations (1.6) and (1.9)), we obtain the fields radiated by the dipole as shown in equation (1.10). In the far field, all components have the structure of transverse electromagnetic (TEM) waves for which the electric and magnetic field amplitudes are related by the free space wave impedance $Z_o \approx 377 \ \Omega$ as shown. Note, for example, that the radiation in the direction $\underline{r} = \underline{p}_0$ is zero (the cross-product is zero), producing the characteristic dumbbell radiation pattern. The radiated magnetic field vector can always be derived from the electric field as shown.

$$\underline{J}(\underline{r}) = i\omega\underline{p}_{0}\delta(\underline{r})$$

$$\Rightarrow \underline{E}(\underline{r}) = \frac{\beta^{2}e^{-i\beta R}}{4\pi\varepsilon_{0}R} \left(\overline{\overline{I}} - \underline{r} \underline{r}\right) \cdot \underline{p}_{0} = \frac{\beta^{2}e^{-i\beta R}}{4\pi\varepsilon_{0}R} \underline{r} \times \left(\underline{r} \times \underline{p}_{0}\right)$$

$$\underline{P}_{0}$$

$$\underline{P}_{0}$$

$$\underline{F}_{0}$$

A magnetic dipole, on the other hand, can be generated by a small loop carrying a uniform current I. The magnetic dipole moment \underline{m} is then defined from the product of current and loop area, and is a vector normal to the plane of the loop,

as shown in equation (1.11).

$$\underline{\underline{H}(\underline{r})} = \frac{\beta^2 e^{-i\beta R}}{4\pi\mu_0 R} \underline{\underline{r}} \times (\underline{r} \times \underline{\underline{m}})$$

$$\underline{\underline{H}(\underline{r})} = -\frac{\beta\omega e^{-i\beta R}}{4\pi R} (\underline{r} \times \underline{\underline{m}})$$
(1.11)

By treating the time variation of loop current I as an equivalent *magnetic* current source in a symmetrized version of Maxwell's equations, the radiated fields can be obtained directly from those of the electric dipole using a duality transformation (Baum, 1995). This symmetry in the equations of (1.1) is useful, as it permits solution of a completely different 'dual' problem to the original without the need for recalculation. Radiation by electric and magnetic dipoles is an example of such dual problems.

The corresponding fields radiated by a magnetic dipole are shown in equation (1.11), where we see that the \underline{E} and \underline{H} fields have been interchanged by the duality transformation, but that the structure of the fields is again due to vector cross and triple products. We shall use these results to formulate scattering by small chiral or handed particles (like a helix), where currents flow in both linear and circular components, in Section 3.3.

1.1.2 Polarised wave coordinate systems

So far our treatment has avoided reference to any specific coordinate system, but in practice the radiation and scattering of waves is projected onto coordinates relevant to the problem at hand. Hence one is faced with the problem of choosing the best coordinate system. One reason why this choice is so important is because we very often want to set up currents \underline{J} on an antenna system so that the radiated wave in the far field has a well-defined orientation or polarisation. However, depending on the coordinates chosen we may find that in some directions the polarisation has components orthogonal to that desired. This is termed crosspolarisation, and in radiation problems is normally undesirable (Collin, 1985). In scattering, on the other hand, it can be useful for identification of the orientation of the induced currents on the scatterer. To illustrate the problems involved in defining crosspolarisation, we outline three commonly used coordinate systems first derived in Ludwig (1973).

1.1.2.1 System I: Cartesian coordinates

This coordinate system is commonly used to describe wave propagation in a paraxial approximation or where there is one well-defined direction. It is defined in terms of a right-handed triplet of unit vectors $\underline{i}, \underline{j}$, and \underline{k} such that the direction of propagation $\underline{k} = \underline{i} \times \underline{j}$, as shown in Figure 1.2. This system can be related to spherical polar coordinates (System II) by transformation equations, as shown in equation (1.12):





Fig. 1.2 Ludwig system I: Cartesian coordinates



Fig. 1.3 Ludwig system II: spherical polar coordinates

For example, consider radiation by an elementary horizontal dipole antenna with dipole moment $\underline{p}_0 = p\underline{i}$. The radiated electric field Cartesian components can then be obtained from equation (1.10) as shown in equation (1.13):

$$\underline{E} = \frac{pe^{-i\beta z}}{4\pi\varepsilon_0 z} \underline{\beta} \times \left(\underline{\beta} \times \underline{i}\right) = \frac{-p\beta^2 e^{-i\beta z}}{4\pi\varepsilon_0 z} \underline{i}$$
(1.13)

This is polarised in the same direction as the antenna current vector. In this way we can consider the EM wave as transferring a 'memory' of the orientation of the dipole source into the far field, with zero crosspolarisation. Such a convenient result does not, however, apply in all coordinate systems, as we now demonstrate.

1.1.2.2 System II: spherical polar coordinates

In theoretical considerations of the radiation and scattering of waves in threedimensional space, spherical polar coordinates are widely used. Here we can locate a source or scatterer at the origin and consider the fields in the surrounding three-dimensional space, as shown in Figure 1.3. The wave propagation direction is then associated with the <u>r</u> unit vector, and the transverse plane formed by the $\underline{\theta}$ and $\underline{\phi}$ unit vectors generates the plane of polarisation of the wave. Figure 1.3 shows how these two unit vectors can be specified by two angles and related to a local Cartesian system. Again considering an elemental x-directed dipole at the origin, we now obtain the radiated field components as shown in equation (1.14):

$$\underline{E} = \frac{p\beta^2 e^{-i\beta R}}{4\pi\varepsilon_0 R} \underline{r} \times (\underline{r} \times \underline{i}) = \frac{p\beta^2 e^{-i\beta R}}{4\pi\varepsilon_0 R} \left(\cos\theta\cos\phi\underline{\theta} - \sin\phi\underline{\phi}\right) \quad (1.14)$$

Here we see that although our source has a well-defined orientation, the radiated field has components that vary with direction and hence are not so neatly constrained as in the Cartesian case. Although providing a convenient general format for three-dimensional radiated fields, the spherical polar system is not the only choice for describing general polarised systems. An alternative, favoured in the antenna measurement community, is based on a hybrid combination of Cartesian and polar concepts, considered as follows.

1.1.2.3 System III: hybrid measurement system

Although the Cartesian and spherical polar systems are convenient for theoretical analyses, in practice antenna patterns and scattering diagrams are referenced to a third coordinate system formed as a hybrid of these two. The key idea here is to define the polarisation unit vectors as Cartesian components <u>i</u>



Fig. 1.4 Ludwig system III: hybrid measurements coordinates

and j, but then to permit three-dimensional field structures by allowing parallel transport of these unit vectors according to spherical polar angles (with the source antenna or scatterer located at the origin). Figure 1.4 shows a schematic of this system. It is clear from the geometry of this transport process that the unit vectors \underline{a}_x and \underline{a}_y are generated by the spherical angle ϕ , as shown in Figure 1.4.

Returning to our example of radiation by an x-directed dipole, we can now establish a systematic method for calculating the level of crosspolarisation radiated by projecting the field in spherical polar coordinates onto the $\underline{a}_x \underline{a}_y$ system. The desired copolarised field is then by definition the a_x component, while a_y is the crosspolarised field. By direct calculation we have the following results:

$$copolarfield = \underline{E} \cdot \underline{a}_{x} = \frac{p\beta^{2}e^{-i\beta R}}{4\pi\varepsilon_{0}R} (\cos\theta\cos^{2}\phi + \sin^{2}\phi)$$

$$crosspolarfield = \underline{E} \cdot \underline{a}_{y} = \frac{p\beta^{2}e^{-i\beta R}}{4\pi\varepsilon_{0}R} \sin\phi\cos\phi(\cos\theta - 1)$$

$$(1.15)$$

Note that in the principal planes (when θ and ϕ are zero) there is zero crosspolarisation. However, for radiation in other directions the ratio of cross- to copolar fields (the XPOL ratio) is given by equation (1.16), which can rise to a maximum of -15 dB when $\phi = \theta = \pi/4$.

$$XPOL = 20 \log_{10} \left(\left| \frac{\sin \phi \cos \phi (\cos \theta - 1)}{\cos \theta \cos^2 \phi + \sin^2 \phi} \right| \right)$$
(1.16)

This level is often too high for radar and communication applications, and hence more sophisticated antennas with even lower crosspolarisation have been developed. To illustrate how such a low crosspolar antenna might be constructed, consider the case of radiation by a Huygens source (Collin, 1985). This can be considered a 'patch' of a plane wave. According to Huygens' principle, such a patch radiates elementary secondary wavelets, the superposition of which marks the advance of the wave front. Figure 1.5 shows such a patch of plane wave of square dimension 2a, where the fields are constant across the aperture and zero elsewhere.

The field radiated by such a structure can be obtained from Maxwell's equations by employing equivalent electric and magnetic currents \underline{J}_{es} , \underline{J}_{ms} in the aperture (Collin, 1985; Cloude, 1995a). These are defined from the transverse components of the field, as shown in equation (1.5). The radiation is then defined by the expression shown. Note that with a distributed current source such as



$$\begin{split} \underline{J}_{es} &= \underline{n} \times \underline{H} \quad \underline{J}_{ms} = \underline{n} \times \underline{E} \\ \underline{E}_{s} &= \frac{e^{-i\beta R}}{R} (1 + \cos\theta) f \left(\cos\phi \underline{\theta} - \sin\phi \underline{\phi} \right) \\ f &= \frac{\sin(\beta \sin\theta \cos\phi a)}{\beta \sin\theta \cos\phi a} \frac{\sin(\beta \sin\theta \sin\phi a)}{\beta \sin\theta \sin\phi a} \end{split}$$

Huygens Source

Fig. 1.5 Radiation by a Huygens patch: the ideal zero cross-polarised source

this, the radiation integral can be explicitly evaluated and produces a Fourier Transform relation between the aperture distribution and the far field. In this case the rectangular distribution gives rise to a SINC function. However, for our purposes, interest centres more on the polarisation properties of the radiated field.

From the polarisation point of view we observe a very interesting result. The radiation from this 'aperture antenna' has zero crosspolarisation in all directions. This shows that in theory, low crosspolarisation can be obtained, although in practice securing the right kind of symmetric aperture distribution can be difficult to engineer, especially over a broad band of frequencies (Collin, 1985; Mott, 1992).

Having established the influence of coordinate systems on the definition of co- and crosspolarised waves in free space, we now turn to consider the propagation of waves in more complex environments. In particular, we consider constraints posed by the presence of the medium on the allowed polarisation states of the propagating field, and thus establish a calculus for dealing with the distortion of the 'memory' effect in the transfer of orientation information from source to far field.

1.2 The propagation of polarised waves

In the absence of sources, waves propagate according to an homogeneous form of Maxwell's equations, as shown on the left in equation (1.17). Complexity now arises in the way in which the presence of material matter influences the way in which the wave can propagate. In this section we consider unbounded wave propagation in each of three special cases: isotropic, when ε and μ are scalar quantities; anisotropic, when $\overline{\varepsilon}$ becomes a tensor or matrix; and chiral materials, where electric and magnetic effects are coupled in the material by helical current flow (Kong, 1985). Without loss of generality we employ the Cartesian coordinate system I with propagation in the +z-direction.

$$\nabla \times \underline{E} = -\frac{\partial \underline{B}}{\partial t} \\ \nabla \times \underline{H} = \frac{\partial \underline{D}}{\partial t} \\ \end{bmatrix} \xrightarrow{\underline{D}} = \varepsilon_r \varepsilon_0 \underline{E} \qquad \underline{D} = \overline{\varepsilon} \underline{E} \qquad \underline{D} = \varepsilon \underline{E} + \eta \underline{B} \\ \xrightarrow{\underline{D}} = \varepsilon_r \varepsilon_0 \underline{E} \qquad \underline{D} = \overline{\varepsilon} \underline{E} \qquad \underline{D} = \varepsilon \underline{E} + \eta \underline{B} \\ \xrightarrow{\underline{B}} = \mu_0 \underline{H} \qquad \underline{B} = \mu_0 \underline{H} \qquad \underline{H} = \gamma \underline{E} + \mu_0^{-1} \underline{B} \\ \xrightarrow{\underline{B}} \text{ Isotropic material Anisotropic material Chiral material}$$
(1.17)

As when considering the radiated field, we first generate a set of vector homogeneous wave equations from the Maxwell curl equations. The resulting systems for each of the three classes of material are shown in equation (1.18).

$$\nabla \times \nabla \times \underline{E} + \varepsilon \mu_0 \frac{\partial^2 \underline{E}}{\partial t^2} = 0 \quad \text{Case I}$$
$$\nabla \times \nabla \times \underline{E} + \overline{\varepsilon} \mu_0 \frac{\partial^2 \underline{E}}{\partial t^2} = 0 \quad \text{Case II} \qquad (1.18)$$
$$\nabla \times \nabla \times \underline{E} - \mu_0 \left(\eta + \gamma\right) \frac{\partial}{\partial t} \nabla \times \underline{E} + \varepsilon \mu_0 \frac{\partial^2 \underline{E}}{\partial t^2} = 0 \quad \text{Case III}$$

We set $\mu = \mu_0$ for applications of interest in remote sensing (That is, we ignore variations in magnetic properties of materials). Also, we postulate vector plane wave solutions propagating in the +z direction of the general form shown in equation (1.19):

$$\underline{E} = e^{i(\omega t - \beta z)} (e_x \underline{i} + e_y j + e_z \underline{k}) = \underline{E}_0 e^{i(\omega t - \beta z)}$$
(1.19)

With these two assumptions we can simplify the homogeneous wave equations to the general form shown in equation (1.20):

$$\nabla \times \nabla \times \underline{E} - \omega^2 \varepsilon \mu_0 \underline{E} = 0 \quad \text{Case I}$$

$$\nabla \times \nabla \times \underline{E} - \omega^2 \overline{\overline{\varepsilon}} \mu_0 \underline{E} = 0 \quad \text{Case II} \quad (1.20)$$

$$\nabla \times \nabla \times \underline{E} - i\omega \mu_0 \left(\eta + \gamma\right) \nabla \times \underline{E} - \omega^2 \varepsilon \mu_0 \underline{E} = 0 \quad \text{Case III}$$

We now seek conditions on the three complex coefficients e_x , e_y and e_z such that the plane wave satisfies the vector wave equations in equation (1.20). To do this we shall make use of the following spatial derivatives of the plane wave solution in our search for a match:

$$\nabla \times \underline{E} = e^{i(\omega t - \beta z)} i\beta(e_y \, \underline{i} - e_x \, \underline{j})$$

$$\nabla \times \nabla \times \underline{E} = e^{i(\omega t - \beta z)} \beta^2(e_x \, \underline{i} + e_y \underline{j})$$
(1.21)

1.2.1 Case I: wave propagation in isotropic media and C2 symmetry

It is one of the unfortunate ambiguities of scientific notation that the word 'polarisation' is used to describe both the electric field orientation of plane waves and also the effect of electric fields on matter. In an attempt to avoid this ambiguity we establish a notation to spell polarisation with 's' to describe wave properties and with 'z' to describe material interactions. Material therefore becomes polarized, while a wave is polarised.

In the simplest case, material becomes polarized by the scalar amplitude of an electric field, and the influence of the material on the wave is then determined by the dielectric constant ε , which in general is a complex scalar. Under these circumstances the wave equation also becomes simplified, and has the form of

a vector Helmholtz equation, as shown in equation (1.22):

$$\underline{D} = \varepsilon \underline{E} \to \nabla \times \nabla \times \underline{E} - \omega^2 \mu \varepsilon \underline{E} = 0$$
(1.22)

In order for the plane wave to be a solution, its components must then satisfy the following equation set obtained by explicit evaluation of equation (1.22):

$$\beta^{2} \begin{pmatrix} e_{x} \\ e_{y} \\ 0 \end{pmatrix} - \omega^{2} \varepsilon \mu_{0} \begin{pmatrix} e_{x} \\ e_{y} \\ e_{z} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \Rightarrow \beta^{2} = \omega^{2} \varepsilon \mu_{0} = \beta_{0} n \Rightarrow$$
$$f = \frac{1}{\sqrt{\varepsilon_{r}} \sqrt{\varepsilon_{0} \mu_{0} \lambda}} = \frac{c}{\sqrt{\varepsilon_{r} \lambda}} = \frac{c}{n\lambda} \qquad (1.23)$$

Note that $e_z = 0$; that is, these plane wave solutions represent transverse electromagnetic (TEM) waves. These waves are also non-dispersive; that is, they all propagate with the same phase velocity, itself determined from the free space velocity $c = 2.997 \times 10^8$ m/s, and the refractive index *n* of the medium, which is related to the square root of the dielectric constant ε_r , as shown in equation (1.23).

These constraints do not specify e_x and e_y . In fact, *any* complex pair will satisfy the wave equation. This we call a C2 symmetry, in that any element of a two-dimensional complex space is a solution. Note, however, that the pair (e_x, e_y) are independent of time and space, and therefore represent a spatio-temporal invariant of the wave. They define the polarisation of the plane wave. Since their resultant always lies in the xy plane transverse to the propagation direction, this is now called the *plane of polarisation* of the wave.

Without loss of generality we can write the pair as a column vector in C2 the space of two-dimensional complex numbers—as shown in equation (1.24), where *m* is the amplitude of the wave, and the trigonometric factors arise directly from the requirement that \underline{w} itself has unit amplitude, or is unitary.

$$\underline{E}_{0} = \begin{bmatrix} e_{x} \\ e_{y} \end{bmatrix} = \sqrt{|e_{x}|^{2} + |e_{y}|^{2}} \begin{bmatrix} \cos \alpha_{w} e^{i\phi_{x}} \\ \sin \alpha_{w} e^{i\phi_{y}} \end{bmatrix} = m \begin{bmatrix} \cos \alpha_{w} e^{i\phi_{x}} \\ \sin \alpha_{w} e^{i\phi_{y}} \end{bmatrix} = \underline{m}\underline{W}$$
(1.24)

We often wish to compare waves with the same amplitude, and therefore set m = 1. In this case the column vector is unitary and has three free parameters. Importantly, each choice of unitary vector \underline{w} then defines a new class of vectors \underline{w}_{\perp} , being orthogonal to the first. As is conventional for complex vector spaces, orthogonality is based on the Hermitian inner product of column vectors, as shown in equation (1.25):

$$\underline{w} = \begin{bmatrix} \cos \alpha_w e^{i\phi_x} \\ \sin \alpha_w e^{i\phi_y} \end{bmatrix} \Rightarrow \underline{w}_{\perp}^{*T} \cdot \underline{w} = 0 \Rightarrow \underline{w}_{\perp} = e^{i\chi} \begin{bmatrix} -\sin \alpha_w e^{i\phi_x} \\ \cos \alpha_w e^{i\phi_y} \end{bmatrix} \Rightarrow \alpha_{\perp} = \alpha_w + \frac{\pi}{2}$$
(1.25)

We see that the orthogonal state is not uniquely defined. There is a phase angle χ left undetermined from <u>w</u> by the combined Hermitian and unitary constraints. This problem can be resolved by considering how the pair <u>w</u> and <u>w</u>_⊥ are to be combined to provide a coordinate system or *polarisation basis* or *frame* for the representation of arbitrary wave states.

To find the components of an arbitrary vector \underline{E} in terms of the unitary states \underline{w} and \underline{w}_{\perp} we form a 2 × 2 transformation matrix through projections, with the unitary vectors as columns, as shown in equation (1.26):

$$\underline{E}' = \begin{bmatrix} \underline{w} & \underline{w}_{\perp} \end{bmatrix} \cdot \underline{E} = \begin{bmatrix} \cos \alpha_w e^{i\phi_x} & -\sin \alpha_w e^{i(\phi_x + \chi)} \\ \sin \alpha_w e^{i\phi_y} & \cos \alpha_w e^{i(\phi_y + \chi)} \end{bmatrix} \cdot \underline{E} = \begin{bmatrix} U \end{bmatrix} \cdot \underline{E} \quad (1.26)$$

We must still deal with the free parameter χ . One way to resolve this issue is to force the matrix U to be special unitary; that is, to have unit determinant. This not only establishes a consistent method for change of base but, as shown in Appendix 2, links directly via group theory to the geometry of the real space of the Poincaré sphere and Stokes vector. With this added condition we obtain the following constraint equation for χ :

$$Det(U) = 1 \Rightarrow \phi_x + \phi_y + \chi = 0 \Rightarrow \chi = -(\phi_x + \phi_y)$$
 (1.27)

Consequently the general *special* unitary change of base matrix can be written as shown in equation (1.28):

$$[U_2] = \begin{bmatrix} \cos \alpha_w e^{i\phi_x} & -\sin \alpha_w e^{-i\phi_y} \\ \sin \alpha_w e^{i\phi_y} & \cos \alpha_w e^{-i\phi_x} \end{bmatrix}$$
(1.28)

Hence we can summarize by saying that if we find a solution to the wave equation \underline{E} in isotropic material, then there is an infinite set of other solutions generated by the relation $[U_2]\underline{E}$. This is a formal representation of the C2 freedom we spoke of in equation (1.23). We see that the properties of special unitary matrices are central to the development of polarimetry theory, and a general description of the properties of such complex matrices is given in Appendix 2. We shall make extensive use of this 2×2 change of base matrix, and also higher-dimensional unitary forms, in analytical manipulations involving polarised waves. To develop $[U_2]$ we involved the idea of orthogonality of complex vectors. In this case it was a mathematical convenience in order to develop a frame or coordinate system. However, orthogonality also arises naturally in many physical systems, as we now consider.

1.2.2 Case II: wave propagation in anisotropic media

In this more complicated case the orientation of the induced polarization vector inside the material is no longer parallel to the orientation of the field excitation, and ε therefore becomes a tensor or matrix. In this case the vector wave equation assumes a tensor form shown in equation (1.29):

$$\frac{\underline{D} = \overline{\overline{\varepsilon}}.\underline{\underline{E}}}{\underline{B} = \mu_0 \underline{\underline{H}}} \right\} \quad \rightarrow \quad \nabla \times \nabla \times \underline{\underline{E}} - \omega^2 \mu_0.\overline{\overline{\varepsilon}} \cdot \underline{\underline{E}} = 0 \tag{1.29}$$

From energy conservation, $\overline{\overline{e}}$ must be a positive definite (PD) Hermitian tensor (see Appendix 1), which means that it is always possible to find a coordinate system inside the material for which the matrix is diagonal (Kong, 1985) and

of the form shown in equation (1.30):

$$\overline{\overline{\varepsilon}} = \begin{bmatrix} \varepsilon_a & 0 & 0\\ 0 & \varepsilon_b & 0\\ 0 & 0 & \varepsilon_c \end{bmatrix} \qquad 0 < \varepsilon_c \le \varepsilon_b \le \varepsilon_a \tag{1.30}$$

Mathematically this is an example of an eigenvalue decomposition, which as we shall see throughout this book often simplifies the treatment of propagation and scattering of polarised waves. As the permittivity tensor is PD Hermitian it has positive real eigenvalues (ε_a , ε_b , ε_c) and orthogonal eigenvectors, which define the *abc* axes of the material. If two of the eigenvalues are equal then the material is *uniaxial*, while if all three are distinct then it is *biaxial*. Such degeneracy can arise through symmetry, as for example in crystal optics, in which cubic symmetry gives rise to triple degeneracy and isotropic propagation. Double degeneracy is found in three crystal groups (tetragonal, hexagonal, and rhombohedral) which are consequently uniaxial. Again we shall see this theme arise in more general scattering problems, whereby symmetry in the medium controls the distribution of eigenvalues of a polarisation matrix.

The *abc* coordinate system forms what are called the *principal axes* of the material, and in general these will not coincide with the xyz of our plane wave propagation system. However, when they do, analysis of propagation greatly simplifies, as we now show. In order for our plane wave to be a solution of the wave equation, the coefficients e_x and e_y must now satisfy the following matrix equation:

$$\beta^2 \begin{pmatrix} e_x \\ e_y \\ 0 \end{pmatrix} - \omega^2 \mu_0 \begin{bmatrix} \varepsilon_{11} & \varepsilon_{12} & \varepsilon_{13} \\ \varepsilon_{21} & \varepsilon_{22} & \varepsilon_{23} \\ \varepsilon_{31} & \varepsilon_{32} & \varepsilon_{33} \end{bmatrix} \cdot \begin{pmatrix} e_x \\ e_y \\ e_z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$
(1.31)

This is generally made complicated because the $\overline{\overline{e}}$ tensor is full. In this case it is more convenient to rewrite equation (1.31) in terms of the electric displacement vector \underline{D} rather than \underline{E} . We then obtain the modified form shown in equation (1.32):

4

$$\nabla \times \nabla \times \begin{bmatrix} \varepsilon_{11} & \varepsilon_{12} & \varepsilon_{13} \\ \varepsilon_{21} & \varepsilon_{22} & \varepsilon_{23} \\ \varepsilon_{31} & \varepsilon_{32} & \varepsilon_{33} \end{bmatrix}^{-1} \cdot \begin{pmatrix} d_x \\ d_y \\ d_z \end{pmatrix} - \omega^2 \mu_0 \cdot \begin{pmatrix} d_x \\ d_y \\ d_z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$
(1.32)

For plane wave solutions, the vector on the far left of this expression has only x and y components, from which it follows that $d_z = 0$; that is, that the <u>D</u> vector (not the <u>E</u> vector) is always transverse to the direction of propagation. For this reason <u>D</u> is often preferred to the electric field <u>E</u> when describing the polarisation of waves in anisotropic media. Now assuming that our external wave system xyz corresponds to *abc* we obtain the following simplified dispersion relation:

$$\beta^2 \begin{bmatrix} \frac{1}{\varepsilon_a} & 0\\ 0 & \frac{1}{\varepsilon_b} \end{bmatrix} \begin{pmatrix} d_x\\ d_y \end{pmatrix} - \omega^2 \mu_0 \cdot \begin{pmatrix} d_x\\ d_y \end{pmatrix} = \begin{pmatrix} 0\\ 0 \end{pmatrix}$$
(1.33)

We see that in this case we no longer have the C2 freedom of isotropic material, and that for a wave to propagate it must be polarised along the a (x) or b (y) directions. Furthermore, the velocity of propagation is different for the two waves—a phenomenon that leads to differential phase shifts between components of the wave, and is known as *birefringence*. Any general polarisation state can be expressed as a linear mixture of a and b through the basis projection matrix of equation (1.28). Thus, when a polarisation state is launched at z = 0 then its a and b components will propagate at different velocities (and also in general with different extinction rates), and hence as it progresses into the material it will *change* its polarisation state. The only exceptions to this are the states a and b themselves. If they are launched into the material then they progress without distortion. If we represent the effect of propagation up to a plane $z = z_0$ as a 2×2 complex matrix $[M_{z0}]$ we can write the following eigenvalue problem:

$$E(z_0) = \begin{bmatrix} M_{z_0} \end{bmatrix} \underline{E}(0) = \lambda \underline{E}(0) \implies (\begin{bmatrix} M_{z_0} \end{bmatrix} - \lambda \begin{bmatrix} I_2 \end{bmatrix}) \underline{E}(0) = 0 \quad (1.34)$$

We then see that the states that remain unchanged due to propagation are *eigenvectors* of the matrix $[M_{z0}]$. Consequently we refer to these as *eigenpropagation states*, or simply *eigenstates*, of the material. We now show how $[M_{z0}]$ can be related to the electric field wave equation.

Returning to equation (1.31) for the electric field, and now imposing the constraint that $d_z = 0$, we can remove the e_z dependence and obtain a pair of equations for e_x and e_y only. The following equation is then obtained for an arbitrary polarisation state, where in the last step we have expressed the spatial term as an ordinary derivative with respect to z, itself obtained from integration of the second derivative appearing from the vector wave equation (assuming $[K_z]$ does not depend on z).

$$\beta^{2} \begin{pmatrix} e_{x} \\ e_{y} \end{pmatrix} - \omega^{2} \mu_{0} \begin{bmatrix} \varepsilon_{11} - \frac{|\varepsilon_{13}|^{2}}{\varepsilon_{33}} & \varepsilon_{12} - \frac{\varepsilon_{13}\varepsilon_{23}}{\varepsilon_{33}} \\ \varepsilon_{12}^{*} - \frac{\varepsilon_{13}^{*}\varepsilon_{23}^{*}}{\varepsilon_{33}} & \varepsilon_{22} - \frac{|\varepsilon_{23}|^{2}}{\varepsilon_{33}} \end{bmatrix} \cdot \begin{pmatrix} e_{x} \\ e_{y} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$
$$\Rightarrow \beta^{2} \underline{E} - \omega^{2} \mu_{0} [K_{z}] \cdot \underline{E} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$
$$\Rightarrow \frac{d^{2} \underline{E}}{dz^{2}} = -\omega^{2} \mu_{0} [K_{z}] \cdot \underline{E}$$
$$\Rightarrow \frac{d\underline{E}}{dz} = [N] \cdot \underline{E} = -i\omega \sqrt{\mu_{0} [K_{z}]} \cdot \underline{E} \qquad (1.35)$$

The most important part of the above analysis is the derivation of a simple matrix differential equation governing the propagation of the C2 column vector \underline{E} in terms of a differential matrix [N], which may be easily integrated to obtain the [M] matrix at distance z_0 , as shown in equation (1.36).

$$\frac{d\underline{E}}{dz} = [N]\underline{E} \Rightarrow \left[M_{z_0}\right] = [M_0] \exp\left(\int_0^{z_0} [N]dz\right)$$
(1.36)

If [N] is constant and we assume $[M_0] = [I_2]$, then this simplifies to equation (1.37):

$$[M_z] = e^{[N]z} (1.37)$$

where the matrix exponential function can be conveniently defined in terms of its infinite series expansion as shown in equation (1.38), which is defined under matrix multiplication for all square matrices [A] (see Appendix 1).

$$\exp([A]z) = I + [A]z + \frac{[A]^2 z^2}{2!} + \dots + \frac{[A]^n z^n}{n!} + \dots$$
(1.38)

We shall now make use of the following six important properties of the matrix exponential function, where the matrix commutator bracket is defined as [A, B] = AB - BA. We see from property II that the eigenvectors of $[M_z]$ and [N] are identical, and that the eigenpolarisation states are determined by the eigenvectors of the reduced dielectric tensor $[K_z]$ in equation (1.35).

$$I = \exp(A) \cdot \exp(B) = \exp(C)$$

$$\Rightarrow C = A + B + \frac{1}{2}[A, B] + \frac{1}{12}([A, [A, B]] + [B, [B, A]]) + \cdots$$

$$II = \exp(SAS^{-1}) = S \exp(A)S^{-1}$$

$$III = \det(\exp(A)) = \exp(Tr(A))$$

$$IV = \exp(A)^{-1} = \exp(-A)$$

$$V = \frac{d}{dz}\exp(Az) = A \exp(Az)$$

$$VI = \frac{d}{dz}\exp(-Az) = -\exp(-Az)A$$
(1.39)

In the special case of zero absorption by the material, $[M_z]$ must be unitary (norm-preserving). If this is the case then its inverse is just its conjugate transpose, and from property IV it follows that [N] = i[H] where [H] is Hermitian. If the matrix $[M_z]$ is special unitary (that is, with unit determinant) then from property III it follows that the matrix [N] must also be traceless. Note that we can always factor a determinant phase term from a unitary propagation matrix $[M_{z0}]$ to leave a special unitary form, as shown in equation (1.40):

$$[M_{z_0}] = [U_2] \cdot \begin{bmatrix} e^{-i\beta_a z_0} & 0\\ 0 & e^{-i\beta_b z_0} \end{bmatrix} \cdot [U_2]^{*T}$$
$$= e^{\frac{-i(\beta_a + \beta_b)}{2} z_0} [U_2] \cdot \begin{bmatrix} e^{-i\frac{(\beta_a - \beta_b)}{2} z_0} & 0\\ 0 & e^{i\frac{(\beta_a - \beta_b)}{2} z_0} \end{bmatrix} \cdot [U_2]^{*T}$$
(1.40)

The determinant phase represents the 'mean' propagation constant in the medium, and the differential terms are all placed inside the special unitary component. We have already encountered special unitary matrices for change of base in C2, and now we see that we can also use them to represent propagation in lossless materials using the matrix exponential function as shown in equation (1.41):

$$[M_2] = [U_2] = \begin{bmatrix} \cos \alpha_w e^{i\phi_x} & -\sin \alpha_w e^{-i\phi_y} \\ \sin \alpha_w e^{i\phi_y} & \cos \alpha_w e^{-i\phi_x} \end{bmatrix} = \exp(iH)$$
$$\Rightarrow [H] = \begin{bmatrix} h_1 & h_2 - ih_3 \\ h_2 + ih_3 & -h_1 \end{bmatrix}$$
(1.41)

where the three coefficients h_1 , h_2 and h_3 are all real. This last result introduces the idea of matrix decomposition to polarimetry. In principle, we take a complex matrix (such as $[M_2]$) and express it as the sum of component matrices, each of which has some simpler physical interpretation. In this way we can 'model' the processes giving rise to the observed matrix in terms of a combination of elementary physical mechanisms. To see this, note that the matrix [H]can be formally expressed as a linear combination of elementary matrices as follows:

$$[H] = \sum_{l=1}^{3} h_l [\sigma_l] = h_1 \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} + h_2 \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} + h_3 \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}$$
(1.42)

The triplet of matrices $\underline{\sigma} = [\sigma_1, \sigma_2, \sigma_3]$ are called the Pauli spin matrices, as they were first applied to problems of spin in quantum mechanics by Wolfgang Pauli (1900–1958). More generally, as we shall see, they are useful for decomposing classical vector wave scattering problems involving complex matrix transformations.

Considering each elementary Pauli matrix at a time, we can use the series expansion of equation (1.38) to derive the corresponding unitary matrices. The key stage is to derive the square of the elementary matrix, and we note that for all three Pauli matrices we have the following relation:

$$\sigma_i^2 = \begin{bmatrix} 1 & 0\\ 0 & 1 \end{bmatrix} = \sigma_0 \tag{1.43}$$

where we have defined a new element σ_0 as the 2 × 2 matrix identity. Hence we can generate the mappings shown in equation (1.44) and give each matrix a simple physical interpretation as follows:

- σ_1 : represents birefringence between the eigenstates a and b.
- σ_2 : represents birefringence between eigenstates at $\pm 45^\circ$ to the basis states; that is, a \pm b.
- σ_3 : represents birefringence between quadrature combinations; that is, a \pm ib, which corresponds, as we can see, to a plane rotation—a result we shall use extensively in this book.

$$\begin{split} \exp(i\theta\sigma_{1}) &= \sigma_{0} + i\theta\sigma_{1} - \frac{\theta^{2}\sigma_{1}^{2}}{2!} + \cdots (i)^{n} \frac{\theta^{n}\sigma_{1}^{n}}{n!} + \cdots \\ &= \sigma_{0} \left(1 - \frac{\theta^{2}}{2!} + \cdots\right) + i\sigma_{1} \left(\theta - \frac{\theta^{3}}{3!} + \cdots\right) \\ &= \cos\theta\sigma_{0} + i\sin\theta\sigma_{1} = \begin{bmatrix} \cos\theta + i\sin\theta & 0\\ 0 & \cos\theta - i\sin\theta \end{bmatrix} \\ &= \begin{bmatrix} 1 & 0\\ 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} e^{i\theta} & 0\\ 0 & e^{-i\theta} \end{bmatrix} \cdot \begin{bmatrix} 1 & 0\\ 0 & 1 \end{bmatrix} \\ \exp(i\theta\sigma_{2}) &= \sigma_{0} + i\theta\sigma_{2} - \frac{\theta^{2}\sigma_{2}^{2}}{2!} + \cdots (i)^{n} \frac{\theta^{n}\sigma_{2}^{n}}{n!} + \cdots \\ &= \sigma_{0} \left(1 - \frac{\theta^{2}}{2!} + \cdots\right) + i\sigma_{2} \left(\theta - \frac{\theta^{3}}{3!} + \cdots\right) \\ &= \cos\theta\sigma_{0} + i\sin\theta\sigma_{2} = \begin{bmatrix} \cos\theta & i\sin\theta\\ i\sin\theta & \cos\theta \end{bmatrix} \\ &= \frac{1}{2} \begin{bmatrix} 1 & -1\\ 1 & 1 \end{bmatrix} \cdot \begin{bmatrix} e^{i\theta} & 0\\ 0 & e^{-i\theta} \end{bmatrix} \cdot \begin{bmatrix} 1 & 1\\ -1 & 1 \end{bmatrix} \\ \exp(i\theta\sigma_{3}) &= \sigma_{0} + i\theta\sigma_{3} - \frac{\theta^{2}\sigma_{3}^{2}}{2!} + \cdots (i)^{n} \frac{\theta^{n}\sigma_{3}^{n}}{n!} + \cdots \\ &= \sigma_{0} \left(1 - \frac{\theta^{2}}{2!} + \cdots\right) + i\sigma_{3} \left(\theta - \frac{\theta^{3}}{3!} + \cdots\right) \\ &= \cos\theta\sigma_{0} + i\sin\theta\sigma_{3} = \begin{bmatrix} \cos\theta & \sin\theta\\ -\sin\theta & \cos\theta \end{bmatrix} \\ &= \frac{1}{2} \begin{bmatrix} 1 & i\\ i & 1 \end{bmatrix} \cdot \begin{bmatrix} e^{i\theta} & 0\\ 0 & e^{-i\theta} \end{bmatrix} \cdot \begin{bmatrix} 1 & -i\\ -i & 1 \end{bmatrix} \end{split}$$

In order to generalize this procedure we need to repeat the series expansion using the most general [H] matrix, itself decomposed as a linear combination of the Pauli matrices. This again requires evaluation of the square of the matrix, which can now be written as shown in equation (1.45):

$$[H]^{2} = (h_{1}\sigma_{1} + h_{2}\sigma_{2} + h_{3}\sigma_{3}) . (h_{1}\sigma_{1} + h_{2}\sigma_{2} + h_{3}\sigma_{3})$$
$$= (h_{1}^{2} + h_{2}^{2} + h_{3}^{2})\sigma_{0} = \theta^{2}\sigma_{0}$$
(1.45)

from which we see it is convenient to define the scalar amplitude of the matrix [H] as θ and to normalize the vector of coefficients $\underline{h} = \theta \underline{n}$ where $\underline{n} \cdot \underline{n} = 1$. With this modification the series again simplifies into elementary trigonometric

functions as follows:

$$\exp(i\theta \underline{n}.\underline{\sigma}) = \sigma_0 + i\theta \underline{n}.\underline{\sigma} - \frac{\theta^2(\underline{n}.\underline{\sigma})^2}{2!} + \cdots (i)^n \frac{\theta^n(\underline{n}.\underline{\sigma})^n}{n!} + \cdots$$
$$= \sigma_0 \left(1 - \frac{\theta^2}{2!} \cdots \right) + i\underline{n}.\underline{\sigma} \left(\theta - \frac{\theta^3}{3!} + \cdots \right)$$
$$= \cos\theta\sigma_0 + i\sin\theta\underline{n}.\underline{\sigma}$$
$$= \begin{bmatrix} \cos\theta + i\sin\theta\underline{n}_1 & i\sin\theta(\underline{n}_2 - i\underline{n}_3) \\ i\sin\theta(\underline{n}_2 + i\underline{n}_3) & \cos\theta - i\sin\theta\underline{n}_1 \end{bmatrix}$$
$$= [U_2] \begin{bmatrix} e^{i\theta} & 0 \\ 0 & e^{-i\theta} \end{bmatrix} [U_2]^{*T}$$
(1.46)

This represents the most general special unitary matrix and an alternative parameterization to that used in equation (1.28). We shall see in Section 1.3 that there is a simple geometrical interpretation of both sets of parameters in terms of spherical trigonometry on the Poincaré sphere. From the form of the eigenvalue decomposition we can see that the general unitary matrix represents birefringence between a pair of orthogonal elliptical polarisations. Such a propagation channel is called a *retarder*, and θ is called the retardence of the channel.

1.2.2.1 Radio wave propagation through the ionosphere

As an important use of the [N] matrix formalism, we now consider the propagation of waves through a gyrotropic or handed medium. An important example of this type is radio wave propagation through a part of the atmosphere called the ionosphere (located at an approximate altitude between 50 and 400 km) (Collin, 1985). Due to ionization by the Sun's radiation, this thin part of the atmosphere can be modelled as a cold plasma in the presence of the Earth's magnetic field. In the absence of a DC magnetic field the dielectric constant of an ionized gas at frequency ω can be written (in the absence of collision damping) in terms of the plasma frequency ω_p as shown in equation (1.47) (Ishimaru, 1991, Chapter 8):

$$\varepsilon_r = 1 - \frac{\omega_p^2}{\omega^2} \qquad \omega_P = \sqrt{\frac{N_e e^2}{m\varepsilon_0}}$$
 (1.47)

where N_e is the electron number density in the material (between 10^{10} and 10^{12} m⁻³ for the ionosphere) and e/m is the charge-to-mass ratio for an electron. Such a material, although frequency-dispersive, is isotropic, and therefore does not distort the polarisation of the propagating wave. However, in the presence of an applied DC magnetic field the situation changes. Here we restrict attention to the case where the DC field is applied along the z-direction (along the direction of propagation for our plane wave). In this case the effect of an electric field depends on its polarisation, and the medium becomes gyrotropic with a dielectric tensor of the form shown in equation (1.48) (Ishimaru 1991,

Chapter 8):

$$\overline{\overline{\varepsilon}} = \varepsilon_0 \begin{bmatrix} \varepsilon_a & i\varepsilon_b & 0\\ -i\varepsilon_b & \varepsilon_a & 0\\ 0 & 0 & \varepsilon_r \end{bmatrix} \Rightarrow \begin{cases} \varepsilon_a = 1 - \frac{\omega_p^2}{\omega^2 - \omega_c^2}\\ \varepsilon_b = \frac{-\omega_c \omega_p^2}{\omega(\omega^2 - \omega_c^2)} \end{cases}$$
(1.48)

where ω_c is the cyclotron frequency defined in terms of the applied magnetic field strength and the charge to mass ratio for the electron, as shown in equation (1.49):

$$\omega_c = \frac{eB_0}{m} \tag{1.49}$$

To give a typical example, the Earth's magnetic field strength is around 5×10^{-4} Tesla, which leads to a cyclotron frequency of 1.4 MHz. Considering propagation in the z direction, we can now use equation (1.48) to generate the $2 \times 2 [K_z]$ matrix directly from this tensor, as shown in equation (1.50):

$$[K_{z}] = \varepsilon_{0} \begin{bmatrix} \varepsilon_{a} & i\varepsilon_{b} \\ -i\varepsilon_{b} & \varepsilon_{a} \end{bmatrix}$$
$$= \frac{\varepsilon_{0}}{2} \begin{bmatrix} 1 & i \\ i & 1 \end{bmatrix} \cdot \begin{bmatrix} \varepsilon_{a} - \varepsilon_{B} & 0 \\ 0 & \varepsilon_{a} + \varepsilon_{B} \end{bmatrix} \cdot \begin{bmatrix} 1 & -i \\ -i & 1 \end{bmatrix}$$
(1.50)

where we have also shown the corresponding eigenvector decomposition of $[K_z]$. This decomposition immediately exposes the physical structure of the propagation problem. The eigenpolarisations are identified as left and right circular polarisation. However, these two states propagate with different propagation constants, determined by the eigenvalues of $[K_z]$.

1.2.2.2 Defining the sense of circular polarisation

Before proceeding, we first establish some notation concerning the handedness of circular polarisation. In common with IEEE engineering standards we define the sense of polarisation from the time variation of the electric field vector in a fixed spatial plane. (Note that spatial variation for a fixed time would be equally valid, but confusingly leads to the opposite definitions.) Again by convention, we define the sense by looking in the -z direction; that is, against the direction of propagation. With this established, left-hand circular is defined as clockwise rotation, and right-hand anticlockwise. These give rise to the polarisation vectors shown in Figure 1.6.

Returning to the gyrotropic medium, we see that left-hand circular polarisation is associated with an eigenvalue $\varepsilon_a - \varepsilon_b$ while right-hand circular polarisation is associated with $\varepsilon_a + \varepsilon_b$. We can now calculate the [N] matrix for this medium, as shown in equation (1.51):

$$[N] = -i\omega\sqrt{\mu_0}\sqrt{[K_z]}$$
$$= -i\omega\sqrt{\varepsilon_0\mu_0}\frac{1}{2}\begin{bmatrix}1 & i\\ i & 1\end{bmatrix}\begin{bmatrix}\sqrt{\varepsilon_a - \varepsilon_b} & 0\\ 0 & \sqrt{\varepsilon_a + \varepsilon_b}\end{bmatrix} \cdot \begin{bmatrix}1 & -i\\ -i & 1\end{bmatrix} \quad (1.51)$$



Fig. 1.6 Definition of left- and right-hand circular polarisations

and finally we obtain the propagation matrix $[M_z]$ using the exponential function as shown in equation (1.52):

$$[M_{z}] = \exp([N]z)$$

= $\frac{1}{2} \begin{bmatrix} 1 & i \\ i & 1 \end{bmatrix} \cdot \begin{bmatrix} \exp(-i\beta_{l}z) & 0 \\ 0 & \exp(-i\beta_{r}z) \end{bmatrix} \cdot \begin{bmatrix} 1 & -i \\ -i & 1 \end{bmatrix}$ (1.52)

where the two propagation constants are defined in equation (1.53):

$$\beta_{l} = \beta \sqrt{\varepsilon_{a} - \varepsilon_{b}} = \beta \sqrt{\left(1 - \frac{\omega_{p}^{2}}{\omega(\omega + \omega_{c})}\right)}$$

$$\beta_{r} = \beta \sqrt{\varepsilon_{a} + \varepsilon_{b}} = \beta \sqrt{\left(1 - \frac{\omega_{p}^{2}}{\omega(\omega - \omega_{c})}\right)}$$
(1.53)

We can see that the right circular wave has a resonance when $\omega = \omega_c$. This wave is forcing the electrons to move in their 'natural' direction about the magnetic field (according to the Lorentz force equation $\underline{F} = q(\underline{E} + \underline{v} \times \underline{B})$). For this reason it is called the *extraordinary* wave. The left circular wave, on the other hand, forces the electrons in the opposite direction and therefore shows no resonance. It is termed the *ordinary* wave. Note that when ω is less than some critical frequency ω_1 then β_L becomes imaginary and the ordinary wave does not propagate. The cut-off frequency can be easily obtained from equation (1.53), as shown in equation (1.54):

$$\omega_1 = \sqrt{\omega_p^2 + \frac{\omega_c^2}{4} - \frac{\omega_c}{2}}$$
(1.54)

Importantly, the extraordinary wave *can* propagate at low frequencies when the ordinary wave is below cut-off. Hence low-frequency waves *can* penetrate the ionosphere along lines of the Earth's magnetic field. This is the main mechanism behind the low-frequency *whistler* mode of atmospheric propagation. These results are summarized in Figure 1.7. Here we show typical dielectric constant variation with frequency and polarisation.

We see that the ordinary wave has a relatively simple behaviour with a cut-off frequency of ω_1 . The extraordinary wave shows more complex behaviour, with



Fig. 1.7 Vector propagation modes in gyrotropic media

two branches to its propagation behaviour, one at low frequencies, and one at high. Note that at high frequencies (compared to ω_c) the medium becomes isotropic and transparent with $\varepsilon_r = 1$. There is, however, a second important polarisation phenomenon arising from this result: the distortion of linear polarisations as they propagate via Faraday rotation, as we now discuss.

1.2.2.3 Faraday rotation

We now consider a Pauli matrix expansion of [N] and show how it leads naturally to a description of Faraday rotation in gyrotropic media. We first rewrite equation (1.52) for $[M_z]$ as shown in equation (1.55):

$$[M_{z}] = \frac{1}{2}e^{-i\frac{\beta_{l}+\beta_{r}}{2}z}\begin{bmatrix}1&i\\i&1\end{bmatrix}\cdot\begin{bmatrix}\exp(-i\Delta\beta z)&0\\0&\exp(i\Delta\beta z)\end{bmatrix}\cdot\begin{bmatrix}1&-i\\-i&1\end{bmatrix}$$
$$= e^{-i\frac{\beta_{l}+\beta_{r}}{2}z}\begin{bmatrix}\cos\Delta\beta z & -\sin\Delta\beta z\\\sin\Delta\beta z & \cos\Delta\beta z\end{bmatrix}$$
$$= e^{-i\frac{\beta_{l}+\beta_{r}}{2}z}\begin{bmatrix}\cos\theta_{F} & -\sin\theta_{F}\\\sin\theta_{F} & \cos\theta_{F}\end{bmatrix}$$
$$= e^{-i\frac{\beta_{l}+\beta_{r}}{2}z}e^{-i\theta_{F}\sigma_{3}}$$
(1.55)

Here we have factored the average propagation constant as indicated in equation (1.40), and defined a differential wavenumber between the left- and right-handed waves as $\Delta\beta = (\beta_l - \beta_r)/2$. By expanding the matrix product we obtain a unitary plane rotation matrix as shown. This in turn may be expressed as the matrix exponential of a single Pauli matrix, σ_3 . The result is that incident linear polarisations are rotated through an angle $\theta_f = \Delta\beta z$. This is called Faraday rotation, and arises as a consequence of the circular polarised eigenstates for gyrotropic media (Ishimaru, 1991; Collin, 1985; Bickel and Bates, 1965). Physically we can consider a linearly polarised wave as decomposed into two

counter-propagating circular waves, and as the two circular components propagate with different velocities so they accrue a phase difference. This phase difference yields a rotation of the linear polarisation state. The connection between phase shifts of circular polarisation and rotations of linear polarisation is of fundamental importance in radar polarimetry, and we shall encounter it several times in our analysis. Again we note that the Pauli matrix decomposition provides a natural formalism for identifying the physical consequences of wave propagation in such media.

One interesting property of Faraday rotation is its invariance to the direction of wave propagation. If we now consider a plane wave propagating in the -z direction as a first step, the above formulae remain the same but with -z replacing z. In this case the rotation matrix is apparently transposed, as the Faraday angle changes sign since $\theta_F = \Delta\beta z$. However, the DC magnetic field has a fixed polarity (+z direction), and hence the matrix [K_z] is conjugated for -z propagation (since the off-diagonal elements change sign with B_0 ; see equation (1.48)):

$$[K_z] = [K_{-z}]^* \tag{1.56}$$

Consequently the left and right circular polarisations exchange eigenvalues, and hence both Δk and z change sign. This leaves the sign of the Faraday angle unchanged, as a consequence of which the matrix for -z propagation $[M_{-z}]$ can be written as follows:

$$[M_{-z}] = e^{i\frac{\beta_I + \beta_T}{2}z} \begin{bmatrix} \cos \theta_F & -\sin \theta_F \\ \sin \theta_F & \cos \theta_F \end{bmatrix} = e^{i\frac{\beta_I + \beta_T}{2}z} e^{-i\theta_F \sigma_3}$$
(1.57)

Surprisingly, the rotation is in the same direction as for +z; that is, if the wave first propagates through the medium and is then returned to its starting point then the Faraday rotation is not cancelled but doubled, since

$$[M_{z}][M_{-z}] = \begin{bmatrix} \cos \theta_{F} & -\sin \theta_{F} \\ \sin \theta_{F} & \cos \theta_{F} \end{bmatrix}^{2} = e^{-i\theta_{F}\sigma_{3}}e^{-i\theta_{F}\sigma_{3}}$$
$$= e^{-i2\theta_{F}\sigma_{3}} = \begin{bmatrix} \cos 2\theta_{F} & -\sin 2\theta_{F} \\ \sin 2\theta_{F} & \cos 2\theta_{F} \end{bmatrix}$$
(1.58)

This can be traced to the presence of the DC magnetic field, which has a polarity of its own and causes this lack of reciprocity. This is in contrast to a second type of circularly polarised wave propagation that occurs in many natural media, such as sugar solutions (optical activity) and in manmade chiral materials such as helical microwave dielectric composites. Here again, circular eigenpolarisations are generated, but this time the effect does not double with space reversal and has a fundamentally different physical origin, as we now consider.

1.2.3 Case III: Propagation in chiral media

Returning to the vector wave equations (1.18), we now consider the allowed propagation states in media with coupled electric and magnetic field effects. The simplest example to consider of such a material is a cloud of small helical

particles embedded in a host material. The application of an electric field will then cause polarization of the particles but also magnetization through circulating induced currents, which in turn will generate a magnetic field. Hence the constitutive material equations require a coupling of electric and magnetic field effects. In the general case all coupling terms can be tensors, as an extension of that described in case II (for a fuller treatment see (Lakhtakia, 1989)). However, an important class of systems can be characterized by scalar coupling terms. These chiral media are characterized by the usual scalar permittivity ε and permeability μ , but also by chiral admittance parameters γ and η such that the constitutive equations have the form shown in equation (1.59) (Ablitt, 1999, 2000):

$$\underline{\underline{D}} = \varepsilon \underline{\underline{E}} + \eta \underline{\underline{B}}$$

$$\underline{\underline{H}} = \gamma \underline{\underline{E}} + \mu_0^{-1} \underline{\underline{B}}$$
(1.59)

For simplicity we here consider the case of lossless chiral media where $\eta = \gamma$ and both are purely imaginary, so the constitutive equations have the special form shown in equation (1.60):

$$\underline{\underline{D}} = \varepsilon \underline{\underline{E}} - i\gamma \underline{\underline{B}}$$

$$\underline{\underline{H}} = -i\gamma \underline{\underline{E}} + \mu_0^{-1} \underline{\underline{B}}$$
(1.60)

We now consider the properties of polarised plane wave propagation in such materials. Before proceeding to the wave equation, we note that by using Maxwell's curl equations we can rewrite these relations in the form shown in equation (1.61):

$$\underline{D} = \varepsilon \left(\underline{E} + \Omega \nabla \times \underline{E} \right)$$

$$\underline{B} = \mu_0 \left(\underline{H} + \Omega \nabla \times \underline{H} \right)$$
(1.61)

where Ω is related to the chiral admittance γ as shown in 1.62:

$$\Omega = \frac{\gamma}{\omega\varepsilon} \tag{1.62}$$

These show that \underline{D} not only depends on the local value of \underline{E} at a point in the material, but also on neighbouring values through the local spatial derivative of \underline{E} . This is termed *spatial dispersion*, and is characteristic of this type of material. With this notation established we now return to the vector wave equation for plane waves in such media, and obtain

$$\nabla \times \nabla \times \underline{E} - 2\omega\mu_0\gamma \nabla \times \underline{E} - \omega^2 \varepsilon \mu_0 \underline{E} = 0$$
(1.63)

Performing the spatial derivatives for the plane wave we obtain the following matrix equation for the electric field components:

$$\beta^{2} \begin{pmatrix} e_{x} \\ e_{y} \\ 0 \end{pmatrix} - 2i\beta\omega\mu_{0}\gamma \begin{pmatrix} e_{y} \\ -e_{x} \\ 0 \end{pmatrix} - \omega^{2}\varepsilon\mu_{0} \begin{pmatrix} e_{x} \\ e_{y} \\ e_{z} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$
(1.64)
from which see that, unlike the case for anisotropic material, $e_z = 0$ is always true and so these are TEM waves. We can now obtain the $[K_z]$ matrix by inspection, as shown in equation (1.65):

$$[K_{z}] = \varepsilon \begin{bmatrix} 1 & \frac{2i\beta\gamma}{\varepsilon\omega} \\ -\frac{2i\beta\gamma}{\varepsilon\omega} & 1 \end{bmatrix} = \varepsilon \begin{bmatrix} 1 & i\varepsilon_{b} \\ -i\varepsilon_{b} & 1 \end{bmatrix}$$
$$= \frac{\varepsilon}{2} \begin{bmatrix} 1 & i \\ i & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 - \varepsilon_{b} & 0 \\ 0 & 1 + \varepsilon_{b} \end{bmatrix} \cdot \begin{bmatrix} 1 & -i \\ -i & 1 \end{bmatrix}$$
(1.65)

The eigenvector decomposition again yields left and right circular eigenpolarisations and differential propagation phase (circular birefringence) due to a splitting of the eigenvalues. Note that because of spatial dispersion this matrix is itself a function of the desired unknown wavenumber β . The [N] matrix can be obtained by taking the square root of [K_z], as shown in equation (1.66):

$$[N] = -i\omega\sqrt{\mu_0}\sqrt{[K_z]}$$
$$= -\frac{i\omega\sqrt{\mu_0\varepsilon}}{2} \begin{bmatrix} 1 & i\\ i & 1 \end{bmatrix} \cdot \begin{bmatrix} \sqrt{1-\varepsilon_b} & 0\\ 0 & \sqrt{1+\varepsilon_b} \end{bmatrix} \cdot \begin{bmatrix} 1 & -i\\ -i & 1 \end{bmatrix}$$
(1.66)

Finally, by using the exponential function we obtain the $[M_z]$ matrix for propagation to z, as shown in equation (1.67):

$$[M_z] = \exp([N]z)$$

= $\frac{1}{2} \begin{bmatrix} 1 & i \\ i & 1 \end{bmatrix} \cdot \begin{bmatrix} \exp(-i\beta_l z) & 0 \\ 0 & \exp(-i\beta_r z) \end{bmatrix} \cdot \begin{bmatrix} 1 & -i \\ -i & 1 \end{bmatrix}$ (1.67)

where the two propagation constants are defined from $\beta_0 = \omega \sqrt{\mu_0 \varepsilon}$ as

$$\beta_L = \beta_0 \sqrt{1 - 2\Omega\beta_L} \Rightarrow \beta_L = \beta_0 (-\beta_0 \Omega + \sqrt{1 + \beta_0^2 \Omega^2})$$

$$\beta_R = \beta_0 \sqrt{1 + 2\Omega\beta_R} \Rightarrow \beta_R = \beta_0 (\beta_0 \Omega + \sqrt{1 + \beta_0^2 \Omega^2})$$
(1.68)

The sign of Ω determines the handedness of the medium as follows: Clockwise or d-rotatory material:

 $\Omega > 0$ $\beta_R > \beta_L$ and the phase velocity for RHC is slower than LHC.

Anticlockwise or l-rotatory material:

 $\Omega < 0$ $\beta_R < \beta_L$ and the phase velocity for LHC is slower than RHC.

The parameter $\beta_0 \Omega$ is typically small, lying in the range $10^{-6} < |\beta_0 \Omega| < 10^{-4}$ for natural materials. In practice we can therefore expand the square root and obtain the following simplified relationships:

$$\beta_L \approx \beta_0 (1 - \beta_0 \Omega) \\ \beta_R \approx \beta_0 (1 + \beta_0 \Omega) \\ \end{bmatrix} \Rightarrow \frac{n_L \approx (1 - \beta_0 \Omega)}{n_R \approx (1 + \beta_0 \Omega)} \\ \Rightarrow \beta_0 \Omega \approx \frac{n_R - n_L}{n_R + n_L}$$
(1.69)

where n_r is the refractive index of the material for right-hand circular polarisation. Thus Ω can be obtained experimentally by measuring the refractive index of the material for right and left circularly polarised wave transmission. Alternatively, the chiral parameter can be estimated by measuring the rotation of a linearly polarised wave. This rotation effect is similar in form to that for Faraday rotation, and can be made explicit by using the Pauli matrix representation of [N] which leads to an $[M_z]$ matrix of the form

$$[M_z] = e^{-i\frac{\theta_l + \theta_r}{2}z} \begin{bmatrix} \cos \theta_A & -\sin \theta_A \\ \sin \theta_A & \cos \theta_A \end{bmatrix} = e^{-i\frac{\theta_l + \theta_r}{2}z} e^{-i\theta_A \sigma_3}$$
(1.70)

where

$$\theta_A = \frac{1}{2} (\beta_L - \beta_R) z \approx -\beta_0^2 \Omega z = \frac{\pi}{\lambda} z \left(n_R - n_L \right)$$
(1.71)

This rotation is clockwise when Ω is positive, and anticlockwise when Ω is negative (looking into the source of the wave). To compare different materials, the *specific rotatory power* is defined as θ_A/z . This has a value ranging from around 20 degrees/mm for solids such as quartz, down to 0.4 degrees/mm for liquids such as turpentine.

Note that a key difference between this and Faraday rotation is the behaviour under inversion of space coordinate. In this case, when we consider propagation in the -z direction, θ_A changes sign, and hence this type of rotation is cancelled with propagation back through the medium. We now turn to consider a generalization of these ideas to enable us to model propagation in arbitrary media.

1.2.4 The Jones calculus: homogeneous and inhomogeneous propagation channels

In the last section we established an important matrix differential equation relating the effect of a propagation channel on the polarisation of a plane wave, as shown in equation (1. 72):

$$\frac{d\underline{E}}{dz} = [N]\underline{E} \implies \left[M_{z_0}\right] = [M_0] \exp\left(\int_0^{z_0} [N]dz\right)$$
(1.72)

We further saw that the matrix [N] can be conveniently expanded in terms of the Pauli spin matrices so that it has the following general form:

$$[N] = i[H] = i \sum_{j=0}^{3} h_j \sigma_j = i \begin{bmatrix} h_0 + h_1 & h_2 - ih_3 \\ h_2 + ih_3 & h_0 - h_1 \end{bmatrix}$$
(1.73)

The study of general solutions of this equation was first carried out in optics in Jones (1941, 1948), and has subsequently been termed the Jones calculus. In this section we outline the general structure of this method and detail a classification of different propagation channels based on the eigenvector decomposition of the propagation matrix (Azzam, 1987; Lu, 1994).

To complete our study of wave propagation, we first consider generalization to the case where the propagation channel involves losses and the matrix $[M_z]$

is no longer unitary. The main extension required is to permit the presence of complex Pauli coefficients h_j . The argument of the exponential function then also becomes complex in general. Nonetheless, the series still splits into odd and even components, and so the limits are modified to hyperbolic rather than trigonometric functions. We have then, in the general case, the following mapping for *complex* $\theta = h_0 + \sqrt{h_1^2 + h_2^2 + h_3^2}$:

$$[M_{2}] = \sqrt{|M_{2}|} e^{iH} = \sqrt{|M_{2}|} \exp(\theta \underline{n} \cdot \underline{\sigma})$$

$$\exp(\theta \underline{n} \cdot \underline{\sigma}) = \sigma_{0} + \theta \underline{n} \cdot \underline{\sigma} + \frac{\theta^{2}(\underline{n} \cdot \underline{\sigma})^{2}}{2!} + \cdots + \frac{\theta^{n}(\underline{n} \cdot \underline{\sigma})^{n}}{n!} + \cdots$$

$$= \sigma_{0} \left(1 + \frac{\theta^{2}}{2!} \cdots\right) + \underline{n} \cdot \underline{\sigma} \left(\theta + \frac{\theta^{3}}{3!} + \cdots\right)$$

$$\Rightarrow [M_{2}] = \sqrt{|M_{2}|} (\cosh \theta \sigma_{0} + \sinh \theta \underline{n} \cdot \underline{\sigma})$$

$$= \sqrt{|M_{2}|} \cdot \begin{bmatrix} \cosh \theta + \sinh \theta n_{1} & \sinh \theta (n_{2} - in_{3}) \\ \sinh \theta (n_{2} + in_{3}) & \cosh \theta - \sinh \theta n_{1} \end{bmatrix} (1.74)$$

where we have factored the determinant of $[M_2]$ as a complex scalar. We will later treat the special case when $[M_2]$ has zero determinant. This matrix is no longer unitary, but represents a combination of differential wave attenuation and phase shifts. Interesting special cases arise when the h_j coefficients are either all real or all imaginary. The $[M_2]$ matrices for these two special cases can be written as shown in equation (1.75).

$$\begin{split} \sqrt{|M_2|} \cdot \exp(\theta \underline{n} \cdot \underline{\sigma}) &= \sqrt{|M_2|} \cdot \begin{bmatrix} \cosh \theta + \sinh \theta n_1 & \sinh \theta (n_2 - in_3) \\ \sinh \theta (n_2 + in_3) & \cosh \theta - \sinh \theta n_1 \end{bmatrix} \\ &= e^{-\phi} \cdot [U_2] \cdot \begin{bmatrix} e^{\theta} & 0 \\ 0 & e^{-\theta} \end{bmatrix} \cdot [U_2]^{*T} \\ \sqrt{|M_2|} \cdot \exp(i\theta \underline{n} \cdot \underline{\sigma}) &= \sqrt{|M_2|} \cdot \begin{bmatrix} \cos \theta + i \sin \theta n_1 & i \sin \theta (n_2 - in_3) \\ i \sin \theta (n_2 + in_3) & \cos \theta - i \sin \theta n_1 \end{bmatrix} \\ &= e^{-i\phi} \cdot [U_2] \cdot \begin{bmatrix} e^{i\theta} & 0 \\ 0 & e^{-i\theta} \end{bmatrix} \cdot [U_2]^{*T} \end{split}$$
(1.75)

These two forms are of special importance in practical applications, as they generate two common classes of propagation channel. We first summarize their general structure, and then show an important result that any propagation channel can be decomposed into a cascade of such effects.

1.2.4.1 Diattenuator channels

For *h* pure imaginary, $[M_2]$ has the general form of a 2 × 2 Hermitian matrix (the upper case in equation (1.75)). Such a channel introduces differential attenuation between the (orthogonal) eigenstates defined by the eigenvector matrix $[U_2]$. It is termed a diattenuator channel (Lu, 1996), and is characterized by its

diattenuation D, defined in equation (1.76):

$$D = \tanh 2\theta \qquad 0 \le |D| \le 1 \tag{1.76}$$

As z tends to infinity, D tends to 1, and such a channel will completely extinguish the eigenvector corresponding to the smallest eigenvalue of [N]. An example of such a case is a polarising filter in optics (such as sheet Polaroid), whereby a propagation matrix of the form shown in equation (1.77) is used to polarise a light source into a preferred state:

$$[M_2] = \begin{bmatrix} \lambda_1 & 0\\ 0 & 0 \end{bmatrix} \tag{1.77}$$

In microwave remote sensing such a situation can also arise for propagation through an oriented grid or volume of aligned scatterers, such as occurs in the growth of many aligned agricultural crops such as wheat stalks (see Section 3.5.2).

1.2.4.2 Retarder channels

For h real, $[M_2]$ has the form of a 2 × 2 unitary matrix (lower case in equation (1.75)). Such a channel shows no attenuation, but introduces differential phase shift between orthogonal eigenstates. It is termed a retarder channel, and is characterized by its retardence *R*, defined in equation (1.78):

$$R = \tan 2\theta \qquad 0 \le R \le \pi \tag{1.78}$$

1.2.5 The polar decomposition

Very often a propagation channel is composed of a cascade of several composite channels, as shown schematically in Figure 1.8. One important property of the Jones calculus is that the $[M_2]$ matrix for the overall channel can be decomposed into a product of $[M_2]$ matrices for each channel so that

$$[M_2] = [M_{zN}] \cdot [M_{zN-1}] \cdot \dots \cdot [M_{z1}]$$
(1.79)

Hence, in general, a propagation channel will show both retardence and diattenuation and the $[M_2]$ matrix will be neither pure Hermitian nor unitary. However, by employing a polar decomposition of the $[M_2]$ matrix we can always express a general channel as an equivalent cascade of just two elements: a pure retarder [U] followed by a pure diattenuator [H] (or *vice versa*) (Lu, 1996). In matrix



Fig. 1.8 General propagation channel as a cascade of composite channels

form we can then always write the following decomposition:

$$[M_2] = [U] \cdot [H_R] = [H_L] \cdot [U] \tag{1.80}$$

where the Hermitian diattenuator matrices are uniquely defined from $[M_2]$, as shown in equation (1.81):

$$[H_R]^2 = [M_2]^{*T} [M_2]$$

$$[H_L]^2 = [M_2] \cdot [M_2]^{*T}$$

(1.81)

If $[M_2]$ is non-singular, then likewise the retarder [U] can be determined as shown in equation (1.82):

$$[U] = [M_2] \cdot [H_R]^{-1} = [H_L]^{-1} [M_2]$$
(1.82)

If $[M_2]$ is singular then [U] is not uniquely defined, but this case can be accommodated by employing the singular value decomposition, as shown below. In the most general case we can write $[M_2]$ in terms of a singular value decomposition (see Appendix 1) as a product of unitary matrices [V] and [W] and a diagonal matrix of singular values [D], as shown in equation (1.83):

$$[M_{2}] = [V] \cdot [D] \cdot [W]^{*T} \Rightarrow \begin{cases} [H_{R}] = [W] \cdot [D] \cdot [W]^{*T} \\ [H_{L}] = [V] \cdot [D] \cdot [V]^{*T} \\ [U] = [V] \cdot [W]^{*T} \end{cases}$$
(1.83)

This decomposition emphasizes the importance of retarder and diattenuator channels, but also leads to an important classification according to the eigenvectors of $[M_2]$. We can see from this result that if $[M_2]$ has orthogonal eigenstates, then [V] = [W] and $[H_R] = [H_L]$; that is, the diattenuator is uniquely defined for $[M_2]$ matrices with orthogonal eigenvectors. These are called *homogeneous* propagation channels, as the diattenuation and retardence *D* and *R* depend only on the eigenvalues and not on the eigenvectors (Lu, 1994).

It is clear that the most general form of homogeneous channel is generated by the product of retarder and diattenuator channel matrices, where the only constraint is that the $[M_2]$ matrix has orthogonal eigenvectors. This can be secured by employing a product of matrix exponentials with the same <u>n</u> vector. We then have the following explicit form for the general lossy homogeneous channel:

$$[M_2] = \sqrt{|M_2|} \cdot \exp(\phi \underline{n} \cdot \underline{\sigma}) \exp(i\chi \underline{n} \cdot \underline{\sigma}) = \exp[(\phi + i\chi)\underline{n} \cdot \underline{\sigma}]$$

= $\sqrt{|M_2|} \cdot \begin{bmatrix} \cosh(\phi + i\chi) + \sinh(\phi + i\chi)n_1 & \sinh(\phi + i\chi)(n_2 - in_3) \\ \sinh(\phi + i\chi)(n_2 + in_3) & \cosh(\phi + i\chi) - \sinh(\phi + i\chi)n_1 \end{bmatrix}$
(1.84)

which is the same as the general lossy matrix shown in equation (1.74) with the substitution $\theta = \phi + \chi$. However, this does not exhaust the possibilities in terms of possible types of propagation channel, as we now show.

If $[M_2]$ has non-orthogonal eigenvectors, then [V] and [W] are no longer equal, and $[H_R]$ and $[H_L]$ are therefore distinct diattenuators with the same

eigenvalues but different eigenvectors. These are called *inhomogeneous* propagation channels.

To illustrate how inhomogeneous channels may arise from a simple cascade of homogeneous elements, consider the example of a two-layer channel composed of a rotator (such as arises with Faraday rotation or optical activity) followed by a polariser. The composite $[M_2]$ matrix is then represented as shown in equation (1.85):

$$[M_2] = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \cdot \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} = \begin{bmatrix} \cos\theta & -\sin\theta \\ 0 & 0 \end{bmatrix}$$
(1.85)

The eigenvectors of $[M_2]$ are not orthogonal and are easily calculated from the eigenvalues, as shown in equation 1.86:

$$\lambda_{1} = 0 \Rightarrow \underline{e}_{1} = \begin{bmatrix} \sin \theta \\ \cos \theta \end{bmatrix} \\ \lambda_{2} = \cos \theta \Rightarrow \underline{e}_{2} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \} \Rightarrow \underline{e}_{1}^{*T} \underline{e}_{2} \neq 0$$
(1.86)

These are propagation states that remain unchanged after transmission through the channel. However, as we shall now show using the SVD, they do not have the optimum transmittance. The singular value decomposition is easily obtained by inspection, as follows:

$$[M_2] = [V] \cdot [D] \cdot [W]^{*T} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \cdot \begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix}$$
(1.87)

from which it follows that the polar decomposition yields an equivalent retarder and pair of diattenuators of the form shown in equation (1.88):

$$[U] = \begin{bmatrix} \cos\theta & -\sin\theta\\ \sin\theta & \cos\theta \end{bmatrix} \qquad [H_L] = \begin{bmatrix} 1 & 0\\ 0 & 0 \end{bmatrix} \qquad [H_R] = \begin{bmatrix} \cos^2\theta & -\sin\theta\cos\theta\\ -\sin\theta\cos\theta & \sin^2\theta \end{bmatrix}$$
(1.88)

The transmittance or gain of the propagation channel is defined as a function of polarisation \underline{E} , as shown in equation (1.89):

$$T = \frac{\left| [M_2] \cdot \underline{E} \right|^2}{\left| \underline{E} \right|^2} = \frac{\underline{E}^{*T} [M_2]^{*T} [M_2] \underline{E}}{\underline{E}^{*T} \underline{E}}$$
(1.89)

The *extrema* of this function define the maximum and minimum transmittance of the channel. This is a classical Rayleigh quotient (see Appendix 1) for which the maximum and minimum values are obtained for the eigenvalues of $[M_2]^{*T}[M_2]$. From the above we see that these are just the square of the singular values of $[M_2]$. We further see that we can decompose the matrix product, as shown in equation (1.90):

$$[M_2]^{*T} [M_2] = [W] \cdot [D] \cdot [W]^{*T}$$
(1.90)

and so the maximum (minimum) transmittance is obtained using states (columns of [W]) which are *not* the eigenstates of the channel. In the above example we see that the maximum transmittance is 1, obtained when the incident state is

$$\underline{E}_{in} = \begin{bmatrix} \cos \theta \\ -\sin \theta \end{bmatrix} \stackrel{\max T}{\Rightarrow} \underline{E}_{out} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$
(1.91)

In this way we see that the singular value decomposition is more relevant to channel gain optimization studies than the eigenvalue expansion of $[M_2]$. Note, however, that the price to pay for maximization of transmittance is distortion of polarisation state. For homogeneous systems there is little difference between expansions, but for inhomogeneous propagation channels the SVD or eigenvalue decompositions must be selected only after consideration of the application in mind.

1.2.6 Propagation in stochastic channels

Much of this book is concerned with problems of wave scattering in random media. It is then of interest to consider the effect of volume scattering on polarised wave propagation and in particular to examine the way in which the Jones calculus needs to be modified. In this section we show how the attenuation of a coherent wave in a scattering medium can be estimated using Foldy's approximation (Tsang, 1985). While strictly valid only for low particle concentrations, this method leads to a simple extension of the Jones calculus for a range of important scattering problems.

The basic geometry is shown schematically in Figure 1.9, in which a wave is propagating in the z-direction through a host medium. This medium, we assume, is isotropic, and thus is characterized by a C2 polarisation degeneracy with a scalar propagation constant β_0 . However, inside this medium are located discrete scattering particles, which for the sake of simplicity we assume are ellipsoidal in shape. We also assume that the position and orientation of these scattering particles are random variables with some underlying probability distributions. We wish to have a means of calculating the equivalent $[M_2]$ and [N] matrices for this channel using properties of the particles themselves. To begin, we note that each particle acts to transform the incident plane wave into a spherical wave (scattering). In addition, each particle can transform the polarisation of the plane wave due to its own shape, orientation and material structure. To account for these two effects we define for each particle a 2 × 2 scattering matrix as shown in equation (1.92)

$$\begin{bmatrix} E_x^{scat} \\ E_y^{scat} \end{bmatrix} = \frac{e^{i(\omega t - \beta r)}}{r} \begin{bmatrix} S_{xx} & S_{xy} \\ S_{yx} & S_{yy} \end{bmatrix} \cdot \begin{bmatrix} E_x^{inc} \\ E_y^{inc} \end{bmatrix} = \frac{e^{i(\omega t - \beta r)}}{r} [S] \cdot \underline{E}_{inc}$$
(1.92)



Fig. 1.9 Schematic representation of general stochastic propagation channel



Fig. 1.10 Derivation of the propagation matrix in scattering media

At present there is some ambiguity with the coordinates defined for this equation. So far we have dealt only with propagation in the z-direction, but clearly the scattering particle generates a spherical wave, which propagates in three dimensions. We delay discussion of the definition of coordinates for the scattering matrix until Section 1.4.2. At the moment we are interested only in the effect of the particle in the same (z) direction as that of the incident wave. This is termed forward scattering, and later we shall see it as a special case of the general scattering problem.

The matrix [S] contains four complex numbers: two diagonal copolar scattering coefficients which change the amplitude and phase of the wave but maintain its polarisation in the same state as that of the incident plane wave. The off-diagonal terms represent the possibility for crosspolarisation. Here the amplitude and phase of the wave is modified by the particle, which now also scatters radiation into the orthogonal polarisation to that incident. We can use this scattering matrix to derive a Fresnel approximation for the change in electric field with z, and so determine the form of the [N] matrix as follows (van de Hulst, 1981). Figure 1.10 shows a component of the electric field at two positions: f_0 is the field at z while f_1 is the new field at position $z + \Delta z$. We are interested in establishing the change in field $f_1 - f_0$ for small Δz (but still large enough so that $\beta \Delta z \gg 1$, otherwise more rigorous diffraction theory is required). The starting point is to decompose both fields into incident (f_i) and scattered (f_s) field components so that we can write

$$\begin{cases} f_0 = f_i \\ f_1 = f_{1i} + f_s \end{cases} \quad f_i = e^{-i\beta z}$$
(1.93)

where we have allowed f_1 to be modified by the presence of scattering particles. The change in incident field is easily calculated from the exponential function so that

$$\Delta f_i = f_{1i} - f_i = -i\beta . \Delta z. f_i \tag{1.94}$$

It remains only to estimate the scattered field component f_s in order to estimate the total change in the field. This can be obtained from the scattering matrix, as shown in equation (1.95):

$$f_s = \sum \frac{e^{-i\beta r}}{r} . s_{xy} . e_0 \tag{1.95}$$

where the sum must be taken over all the particles contributing to the field at $z + \Delta z$. The factor s_{xy} is one element of the scattering amplitude matrix,

depending on our choice of polarisation channel. We now again face the problem that [S] relates a plane wave to a spherical wave, and so formally we must consider variations in x, y and z. However, we are concentrating on forward scattering only, and so any consideration of x,y variations must be constrained to a paraxial or small angle approximation so that only the forward scattering amplitude is appropriate. To secure this, we replace r in equation (1.95) with its approximation shown on the right-hand side of Figure 1.10. We also consider a large number of particles with density N_0 per unit volume. In this case the summation becomes an integral, and we can rewrite equation (1.95) in the form shown in equation (1.96), where we have included analytical evaluation of the resulting Fresnel integral with infinite limits:

$$f_{s} = e^{-i\beta\Delta z} s_{xy} f_{i} N_{0} \left(\iint \Delta z^{-1} e^{-\frac{i\beta(x^{2}+y^{2})}{2\Delta z}} dx dy \right) dz$$

$$\int_{-\infty}^{\infty} e^{-i\frac{\beta x^{2}}{2\Delta z}} dx = \left(\frac{2\pi\Delta z}{i\beta} \right)^{\frac{1}{2}} \Rightarrow f_{s} = e^{-i\beta\Delta z} s_{xy} f_{i} N_{0} \left(-i\frac{2\pi}{\beta} \right) dz$$
(1.96)

Finally, keeping terms only to first order in dz we obtain

$$f_s \approx -i\frac{2\pi N_o}{\beta} f_i \Delta z s_{xy} \tag{1.97}$$

In practice, the particles contributing to this integral will not have the same orientation and size, and we therefore need one further modification to account for variation over these parameters. Note that changing the size or orientation does not alter the Fresnel integral, but requires only that we now replace the constant factor s_{xy} by a configurational average over distributions of size and shape. So, for example, for ellipsoidal particles with dimensions *a*, *b* and *c* and major-axis orientation defined by three Euler angles *A*, *B* and *C* (Goldstein, 1980) we can write:

$$\langle S \rangle = \int_{0}^{\infty} \int_{0}^{\infty} \int_{0}^{\infty} \int_{0}^{2\pi} \int_{0}^{\pi} \int_{0}^{2\pi} s_{xy} \, p(a, b, c; A, B, C) \, da \, db \, dc \, dA \, dB \, dC \qquad (1.98)$$

where p(...) is the probability density function (PDF) of the distribution; and so finally we have the following expression for the scattered field:

$$f_s \approx -i\frac{2\pi N_o}{\beta} \langle S \rangle f_i \Delta z \tag{1.99}$$

This is the desired result which allows us to write the total change of field for each of the four scattering channels separately as shown in equation (1.100), where we note that the crosspolarised channels contain, by definition, no

contribution from the incident field.

$$\Delta f_{xx} \approx \left(-i\beta - i\frac{2\pi N_0}{\beta} \langle S_{xx} \rangle\right) \Delta z f_0$$

$$\Delta f_{xy} \approx \left(-i\frac{2\pi N_0}{\beta} \langle S_{xy} \rangle\right) \Delta z f_0$$

$$\Delta f_{yx} \approx \left(-i\frac{2\pi N_0}{\beta} \langle S_{yx} \rangle\right) \Delta z f_0$$

$$\Delta f_{yy} \approx \left(-i\beta - i\frac{2\pi N_0}{\beta} \langle S_{yy} \rangle\right) \Delta z f_0$$
(1.100)

1.2.7 The Foldy–Lax equations

Collecting these results into matrix form, and using our notation from previous sections, we can now establish the corresponding [N] matrix for propagation in random media, as shown in equation (1.101) (Tsang, 1985):

$$[N] = \begin{bmatrix} -i\beta - i\frac{2\pi N_0}{\beta} \langle S_{xx} \rangle & -i\frac{2\pi N_0}{\beta} \langle S_{xy} \rangle \\ -i\frac{2\pi N_0}{\beta} \langle S_{yx} \rangle & -i\beta - i\frac{2\pi N_0}{\beta} \langle S_{yy} \rangle \end{bmatrix} \Rightarrow [M_z] = \exp([N]z)$$
(1.101)

This provides a formal connection with the Jones calculus developed in the previous section. Again we have two eigenstates for any such material, and these may or may not be orthogonal, depending now on the distribution of particle size, shape and orientation. In general, the medium will act as a mixed retarder/diattenuator, which changes the polarisation of the incident wave as it progresses into the material. As an example, consider propagation through a cloud of dipoles. We assume all particles have the same size and shape and that all are located in the xy plane, and hence their orientation distribution is controlled by a single parameter function $p(\theta)$. The forward scattering matrix for each particle is then a function of just two parameters of the form shown in equation (1.102):

$$S = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} \cdot \begin{bmatrix} \varepsilon & 0 \\ 0 & 0 \end{bmatrix} \cdot \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} = \varepsilon \begin{bmatrix} \cos^2\theta & -\sin\theta\cos\theta \\ -\sin\theta\cos\theta & \sin^2\theta \end{bmatrix}$$
(1.102)

The configurational average can be written as follows:

$$\langle S \rangle = \varepsilon \begin{bmatrix} \int p(\theta) \cos^2 \theta d\theta & -\int p(\theta) \sin \theta \cos \theta d\theta \\ -\int p(\theta) \sin \theta \cos \theta d\theta & \int p(\theta) \sin^2 \theta d\theta \end{bmatrix}$$
(1.103)

In the special case that the distribution is azimuthally symmetric in the xy plane we have

$$p(\theta) = \frac{1}{2\pi} \Rightarrow \langle S \rangle = \frac{\varepsilon}{2} \begin{bmatrix} 1 & 0\\ 0 & 1 \end{bmatrix}$$
(1.104)

which leads to isotropic propagation with zero diattenuation D and the same extinction rate for all polarisations depending on the particle density, size and dielectric constant. This simple example will be very important when we consider coherent volume scattering in Chapter 7, and demonstrates how it is not only the particle shape but also the orientation distribution that determines the propagation channel in random media problems.

We have seen that wave polarisation can be transformed by propagation through a channel. It is therefore of interest to consider the complete set of possible wave polarisation states with a view to investigating what fraction of all states are generated by a particular transforming channel or system. Such a complete set is geometrically represented by the surface of a sphere called the Poincaré sphere, as we now consider.

1.3 The geometry of polarised waves

Geometry is important for the analysis of polarised wave propagation and scattering problems (Deschamps, 1951; Born and Wolf, 1989; Nye, 1999). It often yields a simplified pictorial representation of the complicated processes involved and, more fundamentally, offers a general procedure for the identification of invariants under transformation of polarisation base. In this section we develop a systematic approach to polarisation geometry that emphasizes two key aspects of the problem: namely, the intimate relationship between complex and real number representations of wave polarisation, and the way in which transformation invariants can be derived for general scattering problems. The former will lead us to consider mappings from complex to real spaces, and ultimately to the Poincaré sphere and Stokes vector. The latter will lead us to consider matrix eigenvalue decompositions for the treatment of general wave scattering and propagation problems.

1.3.1 The polarisation ellipse

Our starting point is the structure of the polarisation ellipse. As the electric field $\underline{E}(\underline{r}, t)$ evolves in three-dimensional space and time it traces out a geometrical structure. If we look at the time variation of this structure in a fixed plane transverse to the direction of propagation and, without loss of generality restrict attention to harmonic plane waves, then this locus is always elliptical. To show this, consider the (real) time components of a general harmonic wave propagating in the z-direction. These can be written in the plane z = 0, as shown in equation (1.105):

$$e_{x} = a_{x} \cos \omega t$$

$$e_{y} = a_{y} \cos \left(\omega \left(t - t_{o} \right) \right) = a_{y} \cos \left(\omega t - \phi \right)$$
(1.105)

where a_x and a_y are the amplitudes of the components, and ϕ is the relative phase shift. We can eliminate time from these equations using standard trigonometric



Fig. 1.11 Geometry of the polarisation ellipse

identities to obtain the equation of an ellipse, as shown in equation (1.106):

$$\frac{e_x \cos \phi}{a_x} - \frac{e_y}{a_y} = \sin \omega t \sin \phi = \sqrt{1 - \frac{e_x^2}{a_x}} \sin \phi$$

$$\Rightarrow \frac{e_x^2}{a_x^2} + \frac{e_y^2}{a_y^2} - 2\frac{e_x e_y}{a_x a_y} \cos \phi - \sin^2 \phi = 0$$
(1.106)

This equation is written in terms of three parameters: a_x , a_y and ϕ . However, a more convenient geometrical representation of the ellipse is in terms of two angles θ and τ , defined as shown in Figure 1.11. Note that the ellipse also has amplitude $a = \sqrt{a_x^2 + a_y^2}$ but initially we set this to unity. This simplifies the geometry, but we shall reintroduce it again later when considering connection to the geometry of the Lorentz transformation. The angle θ is the inclination of the major axis of the ellipse, while τ is called the ellipticity angle and is a measure of the shape of the ellipse. It is zero for the special case of linear polarisation, and has a maximum absolute value of $\pi/4$ for circular polarisation. Note that we define positive and negative ranges for both angles. The variation of θ is evident from a consideration of plane rotations, but the sign of τ is related to the sense of the ellipse.

The 'sense' is a consequence of the fact that although the ellipse is a static object, we are actually dealing with a dynamic process: namely, the time evolution of the electric field. This implies that for any given ellipse we can generate a time variation in either the clockwise or anticlockwise direction. Note that this is only true when we have a well-defined direction of propagation (the z axis). For three-dimensional plane waves we shall have to be more careful in the definition of sense. As shown in Figure 1.6, we then define left-hand polarisations as a clockwise rotation (corresponding to positive τ) and right-hand polarisations as anticlockwise (negative τ).

We can further develop this geometry by considering a complex representation of the ellipse. From Figure 1.11, assuming for the moment that $\theta = 0$ and using the complex exponential function, we can write:

$$\underline{\underline{E}} = \begin{bmatrix} \cos \tau \\ i \sin \tau \end{bmatrix} e^{i\omega t} \tag{1.107}$$

where it is understood that the time domain components can be obtained by forming the real part of this expression. We see that by using a complex vector we obtain a time invariant description of the ellipse. Note that in equation (1.107) the time t = 0 origin lies at the tip of the major axis. In general this origin could be located at any point around the ellipse. To allow for this transformation of time origin in the geometry we rewrite the complex representation in the form shown in equation (1.108):

$$\underline{E} = (\underline{a} + i\underline{b}) e^{i\omega t} \Rightarrow \underline{a}.\underline{b} = 0$$
(1.108)

where \underline{a} and \underline{b} are vectors in the x-y plane lying along the major and minor axes of the polarisation ellipse. For a general time origin we then have the following relation:

$$\underline{E} = (\underline{a} + i\underline{b}) e^{i\omega t} e^{i\phi} = (\underline{p} + i\underline{q}) e^{i\omega t} \Rightarrow \begin{cases} \underline{p} = \underline{a}\cos\phi - \underline{b}\sin\phi\\ \underline{q} = \underline{a}\sin\phi + \underline{b}\cos\phi \end{cases} \Rightarrow \underline{p} \cdot \underline{q} \neq 0$$
(1.109)

where the vectors \underline{p} and \underline{q} again lie in the xy plane, but are not orthogonal and correspond to what are called the conjugate semidiameters of the polarisation ellipse. The relationship between $\underline{a}, \underline{b}$ and p, q is summarized in Figure 1.11.

1.3.1.1 Absolute phase of polarised waves

The absolute phase of the wave can now be defined in terms of the ellipse geometry, as shown in equation (1.110):

$$\phi = \frac{1}{2} \arg(\underline{p}^2 - \underline{q}^2 + 2i\underline{p}.\underline{q}) + n\pi \qquad (1.110)$$

The importance of this result is that it relates absolute phase to spatial vectors \underline{p} and \underline{q} . For example, points of circular polarisation are then defined from singularities of ϕ , for which we have the following geometrical conditions, which corresponds in general to a line given by the intersection of two surfaces in space:

$$\underline{p} \cdot \underline{q} = \underline{p}^2 - \underline{q}^2 = 0 \tag{1.111}$$

These circular polarisation lines are termed C-lines, and are important in a full characterization of fields based on their singularities as used, for example, in the study of catastrophe optics (see Nye (1999) for more details).

1.3.1.2 Polarised waves in three dimensions

A second key advantage of using the conjugate semidiameters is that there then exists a straightforward extension to three-dimensional plane waves. This becomes important in applications when we combine waves from multiple sources and directions. In this more general case we can write the electric field vector as shown in equation (1.112):

$$\underline{\underline{E}}\left(\underline{\underline{r}},t\right) = \left[\underline{\underline{P}}(\underline{r}) + i\underline{\underline{Q}}\left(\underline{\underline{r}}\right)\right]e^{i\omega t}$$
(1.112)

where \underline{P} and \underline{Q} are now three-element vectors. From this, the wave polarisation is still an ellipse but now lying in the plane defined by \underline{P} and \underline{Q} with a normal to the plane $\underline{n} = \underline{P} \times \underline{Q}$. Special cases arise when $\underline{P} \times \underline{Q} = 0$, in which case the polarisation is linear and \underline{n} is not defined. Again this will generally occur along special lines in space called L-lines. Circular polarisation occurs for $\underline{P} \cdot \underline{Q} = \underline{P}^2 - \underline{Q}^2 = 0$, and so there exist C-lines also in the three-dimensional case, although these are in general not the same as those for the two-dimensional case. Note that the direction of wave propagation is not so clearly defined in three dimensions, and some care is required in the definition of sense of elliptical polarisation and absolute phase. Although resort can always be made for general fields to the Poynting vector $\underline{S} = \underline{E} \times \underline{H}$, which is the direction of energy flow, this involves a combination of electric and magnetic field components which is undesirable for polarisation studies. To resolve this we again define the absolute phase from the definition of major and minor axes, as shown in equation (1.113), where $A \cdot B = 0$:

$$\underline{P} + i\underline{Q} = e^{i\phi} \left(\underline{A} + i\underline{B}\right) \tag{1.113}$$

The problem is now to relate the phase of an elliptical wave at one point to the phase at a neighbouring point where the ellipse has a different size, shape and orientation vector \underline{n} . One elegant way to resolve this is to use the Pancharatnam geometric phase (Pancharatnam, 1956), defined from the interferogram between the two complex signals. In this way the phase difference between points is defined from maxima and minima in a (real) intensity pattern. Mathematically it is derived from the Hermitian product, as shown in equation (1.114):

$$\psi = \arg(\underline{E}_1^* \underline{E}_2) \tag{1.114}$$

so that when considering a spatial continuum of polarisation ellipses, the infinitesimal phase change can be written as shown in equation (1.115):

$$\delta\psi = \arg\left\{\underline{E}^*\left(\underline{E} + d\underline{E}\right)\right\} = \frac{Im\left(\underline{E}^* \cdot d\underline{E}\right)}{\left|\underline{E}\right|^2} = -\frac{2\underline{B} \cdot d\underline{A}}{\left|\underline{A}\right|^2 + \left|\underline{B}\right|^2} + d\phi \quad (1.115)$$

where the last term results since $\underline{A} \cdot \underline{B} = 0$ implies that $d(\underline{A} \cdot \underline{B}) = d\underline{A} \cdot \underline{B} + d\underline{B} \cdot \underline{A} = 0$. This then confirms that the change in phase can be decomposed into the gradient of the phase *plus* a contribution from the change in geometry. Note that there still remains a problem with this definition, as ψ is not formally integrable, causing problems when we try to define terms such as $\delta \psi / \delta r$, although for the second term, $\partial \phi / \partial r = \nabla \phi$ is acceptable. Hence there is no suitable unique indicator of propagation direction over the whole field. However, we can define

a phase gradient based on the above decomposition of phase as follows:

$$\underline{t} = -\frac{2AB}{A^2 + B^2}\frac{\partial\varepsilon}{\partial r} + \frac{\partial\phi}{\partial r} = -\frac{2AB}{A^2 + B^2}\underline{t}_{\varepsilon} + \nabla\phi \qquad (1.116)$$

where *A* and *B* are the magnitudes of the vectors, and the angle ε is the angle of rotation of the axes <u>*A*</u>, <u>*B*</u> about <u>*n*</u>, defined as <u>*B*</u> \cdot <u>*dA*</u> = *ABd* ε . The component of <u>*t*</u> parallel to <u>*n*</u> represents the twist of the ellipse about <u>*n*</u>, while the transverse component represents the bend of the axial directions of the ellipse in its own plane. This expression is well behaved at most points in the field, and in particular it can be made to yield sensible definitions even along C-lines, when grad ϕ is singular and lines of linear polarisation when <u>*t*</u> is not defined (see Nye (1999) for more details). This means that we can take <u>*t*</u> as a general definition of the propagation direction and then use it to define handedness at each point in the wave field as the sign of the triple product [*AB t*].

1.3.2 Polarisation geometry for paraxial waves

In many applications of interest the full machinery of this three-dimensional wave field formulation is not required, and two-dimensional fields with a well-defined direction of propagation are considered. This is especially true in remote sensing applications. Returning to these, we now develop several important geometrical properties of these simplified fields.

We start by considering the change of absolute phase of a wave as a special unitary transformation of the complex vector in C2. If we multiply one state by a scalar phase factor $\exp(i\phi)$, then in order to secure a *special* unitary transformation we must apply a conjugate phase change of $\exp(-i\phi)$ to the orthogonal state. In transformation terms, therefore, we cannot change the phase of one polarisation state vector without also considering the effect on its orthogonal partner. In matrix terms we can represent such changes of absolute phase as shown in equation (1.117):

$$\underline{E} = \begin{bmatrix} e^{i\phi} & 0\\ 0 & e^{-i\phi} \end{bmatrix} \cdot \begin{bmatrix} \cos \tau\\ i\sin \tau \end{bmatrix} \phi = \omega t_o$$
(1.117)

This expression then generalizes in a straightforward way if the ellipse orientation θ is not zero. In this more general case, the unitary complex vector that represents an ellipse is as shown in equation (1.118):

$$\underline{E} = \begin{bmatrix} \cos\theta & -\sin\theta\\ \sin\theta & \cos\theta \end{bmatrix} \cdot \begin{bmatrix} \cos\tau\\ i\sin\tau \end{bmatrix} = \begin{bmatrix} \cos\alpha_w\\ \sin\alpha_w e^{i\delta} \end{bmatrix}$$
(1.118)

where we have also shown the general element of C2 as defined in equation (1.24). By direct expansion and use of trigonometric identities we can relate the two sets of parameters, as shown in equation (1.119):

$$\cos 2\alpha_w = \cos 2\theta \cos 2\tau \quad \tan 2\theta = \tan 2\alpha_w \cos \delta$$

$$\tan \delta = \tan 2\tau \csc 2\theta \quad \sin 2\tau = \sin 2\alpha_w \sin \delta$$
 (1.119)

Shortly we shall see a simple geometrical interpretation of this result, but for the moment consider a change of ellipticity of the wave from τ to $\tau + \Delta \tau$. This can be represented by a matrix equation as shown in equation (1.120):

$$\begin{bmatrix} \cos\left(\tau + \Delta\tau\right) \\ i\sin\left(\tau + \Delta\tau\right) \end{bmatrix} = \begin{bmatrix} \cos\Delta\tau & i\sin\Delta\tau \\ i\sin\Delta\tau & \cos\Delta\tau \end{bmatrix} \cdot \begin{bmatrix} \cos\tau \\ i\sin\tau \end{bmatrix}$$
(1.120)

Combining this with the rotation matrix involving θ we can extend the analysis by representing not only unitary vectors in C2, but also special unitary matrix transformations, so that we can transform between any pair of ellipses by first rotating the major axis and then changing the ellipticity. The general compound unitary transformation can then be written as shown in equation (1.121):

$$[U_2] = \begin{bmatrix} e^{-i\phi} & 0\\ 0 & e^{i\phi} \end{bmatrix} \cdot \begin{bmatrix} \cos\tau & -i\sin\tau\\ -i\sin\tau & \cos\tau \end{bmatrix} \cdot \begin{bmatrix} \cos\theta & -\sin\theta\\ \sin\theta & \cos\theta \end{bmatrix}$$
(1.121)

This should be compared with the alternative forms shown in equations (1.28) and (1.46). Note that in this case we have obtained a product decomposition with a strong geometrical interpretation. We can also rewrite equation (1.121) as a product of matrix exponential functions as defined in equation (1.42). The result is shown in equation (1.122):

$$[U_2] = (\cos\phi\sigma_o - i\sin\phi\sigma_1)(\cos\tau\sigma_o - i\sin\tau\sigma_2).(\cos\theta\sigma_o - i\sin\theta\sigma_3)$$

= $\exp(-i\phi\sigma_1)\exp(-i\tau\sigma_2).\exp(-i\theta\sigma_3)$ (1.122)

This demonstrates how we can consider the effect of a general transformation as a compound of three simpler operations, each generated by a single Pauli matrix as shown.

So far our treatment has considered the use of complex matrix transformations for representation of the polarisation ellipse. There is, however, an alternative approach based entirely on real quantities. This has more than theoretical significance. By transforming a 'complex' problem into an equivalent 'real' one, we can dispense with the need for a coherent measurement system. This is of great practical benefit, especially at high frequencies where it is difficult to design detectors capable of following the fast field oscillations directly. Visible optics is a prime example in which widespread use is made of this mapping for polarisation studies. A simple example of this complex-toreal equivalence is in the formation of an interference pattern. Here the spatial variation of a 'complex' field is visualized as a 'real' fringe intensity variation across the field of view. In polarimetry this real-to-complex mapping is more subtle but nonetheless central to the development of techniques for the interpretation of vector scattering problems. In this context it leads to the concept of the Poincaré sphere, as follows.

1.3.3 The Poincaré sphere

We start by considering each of the complex matrix factors in equation (1.122) to show how it can be mapped into an equivalent orthogonal transformation of

a real three-vector \underline{r} . In the context of polarimetry this is called the Stokes vector representation, after the nineteenth-century English mathematician George Gabriel Stokes (1819–1903) (Born and Wolf, 1989; Bickel, 1985; Schmeider, 1969).

The first stage is to define a suitable real vector \underline{r} . As there are three matrices in the cascade of equation (1.122) we anticipate a three-dimensional real space R3 and so set $\underline{r} = (x, y, z)$. The next key idea is to recognize that $|\underline{r}| = \sqrt{x^2 + y^2 + z^2}$ (amplitude) should be a transformation invariant, so that we deal with orthogonal transformations in R3. As we are going to deal with matrix products we need to generate a matrix from \underline{r} with a metric that is a function of the desired invariant. One way to do this is to construct a matrix using the Pauli spin matrices, as shown in equation (1.123):

$$[R] = x\sigma_1 + y\sigma_2 + z\sigma_3 = \begin{bmatrix} x & y - iz \\ y + iz & -x \end{bmatrix} \Rightarrow \det([R]) = -(x^2 + y^2 + z^2)$$
(1.123)

[*R*] Is called a spin matrix, and is constructed so that its determinant is a function of the desired invariant. Note the negative sign. This will be useful when we consider extension to Lorentz transformations in Section 1.5. Having established a 2×2 matrix representation of <u>r</u>, we can now generate transformations using our complex 2×2 exponential functions so that in general we can form a transformation as shown in equation (1.124) (Goldstein, 1980; Misner, 1973):

$$\left[R\left(\varphi,\underline{n}\right)\right] = \exp\left(i\varphi\underline{\sigma}.\underline{n}\right).\left[R\right].\exp\left(-i\varphi\underline{\sigma}.\underline{n}\right)$$
(1.124)

This equation preserves the determinant of [R] on both sides, as the matrix exponentials both have unit determinant and the determinant of a matrix product is just the product of determinants (which was the reason why we focused on special unitary matrices in equation (1.26)). We have therefore constructed a transformation of r with the correct metric.

We must now determine other invariants under this transformation. In particular we seek the axis of any plane rotations. This will allow us to finally determine a mapping from the 2 × 2 complex matrix into a real 3 × 3 orthogonal matrix. To do this we need only consider infinitesimal transformations, so we can conveniently use only the first few terms of the series expansion of the exponential function (see equation (1.38)). Equation (1.125) shows this expansion. We note that since $[R] = \sum_{j=1}^{3} r_j \sigma_j$, invariants can then be conveniently identified from products of the Pauli matrices that commute; that is, for which $\sigma_i \sigma_j - \sigma_j \sigma_i = 0$.

$$[R(\varphi)] = (\sigma_0 + i\varphi\underline{\sigma} \cdot \underline{n} + \cdots) \cdot [R] \cdot (\sigma_0 - i\varphi\underline{\sigma} \cdot \underline{n} + \cdots)$$

$$\approx [R] + i\varphi ([R]\underline{\sigma} \cdot \underline{n} - \underline{\sigma} \cdot \underline{n} [R])$$
(1.125)

Alternatively we can (in this special case of three-dimensional vectors) proceed by noting that the product of two spin matrices can be expressed in terms of scalar and vector products of the original real vectors as $(\underline{a} \cdot \underline{\sigma}) (\underline{b} \cdot \underline{\sigma}) = \underline{a} \cdot \underline{b} + i (\underline{a} \times \underline{b}) \underline{\sigma}$. Hence it follows from this and equation (1.125) that the only other invariant is given by equation (1.126):

$$[R] \underline{\sigma} \cdot \underline{n} - \underline{\sigma} \cdot \underline{n} [R] = \underline{r} \cdot \underline{n}$$
(1.126)

This is a very important result. It shows that the transformation of equation (1.124) has a very simple interpretation in the space of the vector r. Recall that n is a unit real vector but lies in the original complex domain (equation (1.46)), and so this relation demonstrates that under the transformation the component of r in the direction of n is invariant. This is just another way of defining a plane rotation about the *n* axis, but critically it is located in the 'real' space of the vector r. Hence we have derived a mapping from unitary transformations in C2 into orthogonal transformations in R3-a real three-dimensional space. We can now generate this mapping for each of the elementary matrix exponentials in equation (1.44). Using the fact that $\underline{n}_1 = (1, 0, 0), \underline{n}_2 = (0, 0, 1),$ and $\underline{n}_3 = (0, 0, 1)$ we obtain the explicit mappings shown in equation (1.127). Note that double angles appear in the 'real' space. This is a direct consequence of employing the triple matrix product in equation (1.124). It means that the mapping from complex to real is unfortunately not 1-to-1 but 2-to-1. For example, if we add 2π radians to any of the angles in the real space we obtain an unchanged rotation matrix. However, in the corresponding complex space we have two distinct matrices, since the half-angle used in the mapping yields only a π change in the complex domain. So, for every real rotation matrix there are two corresponding unitary partners:

$$\exp(i\theta\sigma_3) = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} \Leftrightarrow \begin{bmatrix} \cos 2\theta & \sin 2\theta & 0 \\ -\sin 2\theta & \cos 2\theta & 0 \\ 0 & 0 & 1 \end{bmatrix}$$
$$\exp(i\tau\sigma_2) = \begin{bmatrix} \cos\tau & i\sin\tau \\ i\sin\tau & \cos\tau \end{bmatrix} \Leftrightarrow \begin{bmatrix} \cos 2\tau & 0 & \sin 2\tau \\ o & 1 & 0 \\ -\sin 2\tau & 0 & \cos 2\tau \end{bmatrix}$$
$$\exp(i\phi\sigma_1) = \begin{bmatrix} \exp(i\phi) & 0 \\ 0 & \exp(-i\phi) \end{bmatrix} \Leftrightarrow \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos 2\phi & \sin 2\phi \\ 0 & -\sin 2\phi & \cos 2\phi \end{bmatrix}$$
$$[U_2] = \exp(i\phi\sigma_1) \exp(i\tau\sigma_2) \exp(i\theta\sigma_3) = \exp(i\chi\underline{\sigma}.\underline{n}) = \cos\chi\sigma_0 + i\sin\chi\underline{\sigma}.\underline{n}$$
$$\Leftrightarrow [O_3] = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos 2\phi & \sin 2\phi \\ 0 & -\sin 2\phi & \cos 2\phi \end{bmatrix} \cdot \begin{bmatrix} \cos 2\tau & 0 & \sin 2\tau \\ o & 1 & 0 \\ -\sin 2\tau & 0 & \cos 2\tau \end{bmatrix} \cdot \begin{bmatrix} \cos 2\theta & \sin 2\theta & 0 \\ -\sin 2\theta & \cos 2\theta & 0 \\ -\sin 2\tau & 0 & \cos 2\tau \end{bmatrix} \cdot \begin{bmatrix} \cos 2\theta & \sin 2\theta & 0 \\ -\sin 2\theta & \cos 2\theta & 0 \\ 0 & 0 & 1 \end{bmatrix}$$
(1.127)

We can formalize this correspondence by defining the groups SU(2) and O_3^+ as, respectively, the group of all 2 × 2 complex unitary and 3 × 3 real orthogonal matrices with determinant +1 (see Appendix 2). If two groups X and Y have a 1-to-1 correspondence between elements they are said to be isomorphic, while if the correspondence is many-to-one they are homomorphic. Hence we have demonstrated that SU(2) and O_3^+ are homomorphic. This SU(2)– O_3^+ homomorphism is fundamental to the development of polarisation geometry. One of



Fig. 1.12 The Poincaré sphere and polarisation space

the most important consequences of this mapping is that matrix multiplication is preserved, and so a cascade of matrices in the complex domain maps into a cascade of matrices in the real domain. This, combined with the result of equation (1.124), allows us to map any unitary matrix as a set of Euler angle transformations in the real space, as shown in equation (1.127). This in turn leads to a simple geometrical interpretation of a unitary matrix transformation as rotations over a sphere in R3. Each point P on the surface of this sphere corresponds to a triplet of angles (θ, τ, ϕ) as shown in Figure 1.12. Here we can see on the right how the angles can be used to locate any point P in the space of the vector r. The two most important angles are θ and τ . These are the longitude and latitude of the point P on the surface of a sphere. The third angle ϕ corresponds to a rotation about the radius of the sphere and, as shown in equation (1.117), relates to the absolute phase of the wave. The set of all polarisations of the same amplitude then covers the surface and generates what is called the Poincaré sphere, after the French mathematician Henri Poincaré (1854–1912). This sphere has all linear polarisations mapped around the equator, with left circular at the north pole and right circular at the south pole. The upper and lower hemispheres then correspond to elliptical states with opposite sense.

One very important consequence of the double angle mapping is that orthogonal polarisations lie diametrically opposite on the sphere and correspond to antipodal points. This is in distinction to the conventional geometrical notion of orthogonality, which involves 90-degree angular separation between vectors. We can also use this sphere to interpret the relationships derived in equation (1.119). The angles θ and τ are then related to α , δ as elements of a spherical triangle, as illustrated on the left-hand side of Figure 1.12. In this way we see them as alternate angular coordinates to locate a polarisation state P in polarisation space.

To proceed, we must now modify this formalism to include the amplitude of the wave, a. This we can do by modifying the spin matrix [R], as shown in equation (1.128) (Misner, 1973; Goldstein, 1980).

$$[R] = a(\sigma_0 + x\sigma_1 + y\sigma_2 + z\sigma_3) = a \begin{bmatrix} 1 + x & y - iz \\ y + iz & 1 - x \end{bmatrix}$$

$$\Rightarrow \det([R]) = a^2(1 - (x^2 + y^2 + z^2))$$
(1.128)

Here we see that by adding the amplitude as a multiple of the identity matrix we can secure a new invariant: namely, the difference between the amplitude squared and the square of the length of the vector \underline{r} . For a polarised wave this metric will always be zero. This modifies the geometry to that of a Lorentz

transformation, where equation (1.128) arises naturally through the use of the Minkowski metric. The Lorentz spin transformation is then obtained as a generalization of equation (1.124) as shown in equation (1.129), where [L] is called a Lorentz spin matrix, det([L]) = 1, $but \gamma$ can be complex, and [L] does not have to be unitary.

$$\begin{bmatrix} R(\gamma, \underline{n}) \end{bmatrix} = \begin{bmatrix} L(\gamma, \underline{n}) \end{bmatrix} \cdot \begin{bmatrix} R \end{bmatrix} \cdot \begin{bmatrix} L(\gamma, \underline{n}) \end{bmatrix}^{*T} \quad \begin{bmatrix} L \end{bmatrix} = \exp(i\gamma \underline{\sigma} \cdot \underline{n}) \quad \gamma = \gamma_r + i\gamma_i$$
(1.129)

Physically this allows us to consider differences in scattering amplitude or absorption across polarisation channels, and therefore represents an important extension of the geometry.

1.3.4 The Stokes vector

We note that by including the amplitude 'a' in a description of polarised waves, we must now formally consider not just a three-vector \underline{r} but a four-vector \underline{g} , called the Stokes vector of the wave, formed as a composite of the three-vector \underline{r} and the wave intensity with elements given by the Pauli matrix expansion of the spin matrix [*R*], as shown in equation (1.130) (Bickel, 1985; Born and Wolf, 1989).

$$\underline{g} = \begin{bmatrix} g_0 \\ g_1 \\ g_2 \\ g_3 \end{bmatrix} = a^2 \begin{bmatrix} 1 \\ \cos 2\theta \cos 2\tau \\ \sin 2\theta \cos 2\tau \\ \sin 2\tau \end{bmatrix}$$
$$= a^2 \begin{bmatrix} 1 \\ \cos 2\alpha_w \\ \sin 2\alpha_w \cos \delta \\ \sin 2\alpha_w \sin \delta \end{bmatrix} \quad g_0^2 - g_1^2 - g_2^2 - g_3^2 = 0 \qquad (1.130)$$

Figure 1.13 shows some examples of the Stokes vectors for linearly and circularly polarised waves. Equation (1.130) defines the Stokes vector in terms of the geometrical parameters of the ellipse. An alternative definition is based on the wave coherency matrix [*J*]. This is a 2×2 Hermitian matrix, and is defined from the product of the complex electric field vector with its conjugate transpose. Using α and δ parameters this matrix can be defined as shown in equation (1.131):

$$\underline{E} = a \begin{bmatrix} \cos \alpha_w \\ \sin \alpha_w e^{i\delta} \end{bmatrix}$$
$$\Rightarrow [J] = \underline{E} \underline{E}^{*T} = a^2 \begin{bmatrix} \cos^2 \alpha_w & \sin \alpha_w \cos \alpha_w e^{-i\delta} \\ \sin \alpha_w \cos \alpha_w e^{i\delta} & \sin^2 \alpha_w \end{bmatrix} = [R]$$
(1.131)

To make this consistent with equation (1.128) the matrix [J] must be related to the Stokes vector g as shown in equation (1.132), where I, Q, U and V are

→	Horizontal Polarisation	g = (1,1,0,0)
Ť	Vertical Polarisation	g = (1,-1,0,0)
/	+45 degree Polarisation	g = (1,0,1,0)
	-45 degree Polarisation	$\underline{g} = (1,0,-1,0)$
\bigcirc	Left Hand circular	g = (1,0,0,1)
\bigcirc	Right Hand circular	g = (1,0,0,-1)



symbols widely used in the optics literature for the Stokes parameters, elements of the Stokes vector.

$$[J] = \begin{bmatrix} E_x E_x^* & E_x E_y^* \\ E_y E_x^* & E_y E_y^* \end{bmatrix} = \frac{1}{2} \begin{bmatrix} g_0 + g_1 & g_2 - ig_3 \\ g_2 + ig_3 & g_0 - g_1 \end{bmatrix} \Rightarrow \underline{g} = \begin{bmatrix} E_x E_x^* + E_y E_y^* \\ E_x E_x^* - E_y E_y^* \\ E_x E_y^* + E_y E_x^* \\ i(E_x E_y^* - E_y E_x^*) \end{bmatrix} = \begin{bmatrix} I \\ Q \\ U \\ V \end{bmatrix}$$
(1.132)

[J] has the following two important properties:

- Det([*J*]) = 0. This follows from the Minkowski metric, and is true for all polarised monochromatic waves.
- $[J] = [J]^{*T}$; that is, [J] is complex Hermitian, and therefore has orthogonal eigenvectors and real eigenvalues (see Appendix 1). For polarised waves it follows that the eigenvector decomposition of [J] can be written as shown in equation (1.133):

$$[J] = \begin{bmatrix} \cos \alpha e^{i\phi} & -\sin \alpha e^{-i\delta} \\ \sin \alpha e^{i\delta} & \cos \alpha e^{-i\phi} \end{bmatrix} \cdot \begin{bmatrix} \lambda_1 & 0 \\ 0 & 0 \end{bmatrix} \cdot \begin{bmatrix} \cos \alpha e^{-i\phi} & \sin \alpha e^{-i\delta} \\ -\sin \alpha e^{i\delta} & \cos \alpha e^{i\phi} \end{bmatrix}$$
(1.133)

In publications not concerned directly with the geometrical structure of the Stokes vector, the modified Stokes vector \underline{g}_m is often used (Tsang, 1985). This again is defined from the wave coherency matrix but in a more straightforward

manner, as shown in equation (1.134):

$$\underline{g}_{M} = \begin{bmatrix} |E_{x}|^{2} \\ |E_{y}|^{2} \\ 2Re(E_{x}E_{y}^{*}) \\ 2Im(E_{x}E_{y}^{*}) \end{bmatrix} \Rightarrow \begin{cases} \underline{g} = [G]\underline{g}_{M} = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \underline{g}_{M} \\ \underline{g}_{M} = [G]^{-1}\underline{g} = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & 0 & 0 \\ \frac{1}{2} & -\frac{1}{2} & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \underline{g}$$
(1.134)

The matrices [G] and $[G]^{-1}$ may then be used to transform between the two conventions. The use of matrices to represent changes in Stokes vectors forms the basis of an important calculus. In equation (1.135) we show two other important examples of matrices used in Stokes algebra. The first $[M]_{conj}$ represents the physical process of conjugation, or taking the complex conjugate of the wave. As shown, this has the effect of changing the sign of the fourth Stokes parameter or changing the sense of the polarisation ellipse. On the right we show the transformation matrix that converts a Stokes vector into its orthogonal polarisation state. This involves reflection of the point in the origin to obtain the coordinates of the antipodal point on the Poincaré sphere.

$$[M]_{conj} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix} \qquad [M]_{\perp} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix}$$
(1.135)

Finally, we note that an extension of this Stokes approach to cover the case of three-dimensional waves, when the wave coherency matrix [J] becomes 3×3 , has been developed by Roman (1959a) as a set of nine Hermitian matrices analogous to the Pauli spin matrices, the coefficients of which define a generalized set of Stokes parameters r_i for the representation of waves with arbitrary form (Roman 1959b). In this case the generalization of equation (1.132) is shown in equation (1.136):

$$\begin{aligned} [J_3] &= \begin{bmatrix} E_x E_x^* & E_x E_y^* & E_x E_z^* \\ E_y E_x^* & E_y E_y^* & E_y E_z^* \\ E_z E_x^* & E_z E_y^* & E_z E_z^* \end{bmatrix} \\ &= \begin{bmatrix} r_0 + r_1 + r_4 & r_2 + ir_3 & r_7 + ir_8 \\ r_2 - ir_3 & r_0 & r_2 - r_5 + i(r_3 - r_6) \\ r_7 - ir_8 & r_2 - r_5 - i(r_3 - r_6) & r_0 - r_1 + r_4 \end{bmatrix} (1.136)$$

Here we concentrate on the simplified 2×2 case, which is the most commonly used in remote sensing. However, special applications that require analysis of dynamic vector near field effects should make use of this more general formalism.

We have now considered the generation and geometry of polarised waves and the distortion effects caused by propagation channels. Now we turn to consider the final stage in Figure 1.1: namely, the scattering of polarised waves and the interpretation of a scatterer as a transformer of polarisation state.

1.4 The scattering of polarised waves

When a plane wave illuminates an object, metallic or dielectric, vector currents are induced, the magnitude and direction of which depend on the shape and material composition of the object. These in turn act as secondary sources that reradiate into space. In general, this reradiation occurs in directions different to that of the incident wave. This is termed wave scattering, and forms the basis for our studies of remote sensing using electromagnetic waves. In particular we are concerned with the polarisation properties of the scattered wave, and to what extent they maintain a 'memory' of the original vector nature of the induced currents.

In order to formulate a general scattering problem we first locate the scatterer in free space at the origin of a spherical polar coordinate system, and then define an incident plane wave direction A using two angles θ and ϕ . This wave has a polarisation with C2 freedom, defined in the transverse xy plane as shown in Figure 1.14. The wave is then scattered, and interest centres on the radiation scattered into a new direction B, again defined by two polar angles θ' and ϕ' . The two wave-direction vectors $\underline{\beta}_i$ and $\underline{\beta}_s$ define a plane called the *scattering plane*, and the bisector between the incident and scattered wave vectors is called the *bisectrix* vector. This bisector plays an important role in a geometrical interpretation of wave scattering, as we show later. It also provides an important symmetry axis for consideration of the basic polarisation properties of the scattering system. The angle in this plane between incident and scattered wave directions is called the scattering angle, and important special cases arise for forward scattering $\psi = 0^\circ$, lateral scattering $\psi = 90^\circ$, and backscatter $\psi = 180^\circ$.

However complicated the scattering problem may be, the scattered fields obey Maxwell's equations, which form a linear system of equations. Hence there is no loss of generality in postulating a linear mapping from A to B. Since both A and B have C2 degeneracy, this mapping can be represented by a 2×2 complex matrix [S] called the amplitude or scattering matrix. In this way an arbitrary polarisation \underline{E}_i at A is mapped by the scatterer into a state \underline{E}_s at B by the relationship shown in equation (1.137), where the factor in front of the matrix accounts for the phase and amplitude variation of a spherical wave of



Fig. 1.14 General wave scattering coordinates for the amplitude scattering matrix

radius 'r' centred on the scattering point:

$$\underline{\underline{E}}_{s} = \frac{e^{-i\beta r}}{r} \begin{bmatrix} S_{xx'} & S_{x'y} \\ S_{y'x} & S_{yy'} \end{bmatrix} \underline{\underline{E}}_{i} = \frac{e^{-i\beta r}}{r} \begin{bmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{bmatrix} \underline{\underline{E}}_{i}$$
(1.137)

Note that sometimes in the literature (van de Hulst, 1981; Hovenier, 2004) a normalizing factor of $1/i\beta$ is used on the right-hand side in equation (1.137). This is supported by the discussion leading up to equation (1.99), but readers should be warned that there exist slight differences of definition for the scattering amplitude matrix.

Cleary, knowledge of this mapping (that is, knowledge of all four complex numbers in [S]) enables full polarimetric characterization of the scattering problem from A to B. We can then determine the amplitude and polarisation of the scattered wave at B for *any* incident polarisation state \underline{E}_i on the Poincaré sphere by combining the C2 freedom with matrix multiplication arising from system linearity. This is called polarisation synthesis or virtual polarisation adaptation (Giuli, 1986; Poelman, 1991). The key idea is that measurement of [S] can lead to adaptive selection of polarisation to enhance features based on signal processing rather than on physical antenna changes. The matrix [S] therefore plays an important role in polarimetry theory (Luneburg, 1996, 1997), and we now turn to discuss its properties in more detail.

1.4.1 The scattering amplitude matrix [S]

A key benefit of knowledge of the [S] matrix is that it provides independence from the measurement basis used. Above, we chose xy and x'y' as a natural coordinate system, but in principle any pair of C2 orthogonal states will be equally acceptable. This can be formally represented by allowing *independent* unitary transformations of the incident and scattered coordinates (Kostinski, 1986; Luneberg, 1996, 1997) so that the change of frame is represented by the pair of matrix equations shown in equation (1.140), where in general $U_{2A} \neq U_{2B}$:

$$\underline{E}'_{i} = [U_{2A}] \underline{E}_{i}$$

$$\underline{E}'_{s} = [U_{2B}] \underline{E}_{s}$$
(1.140)

With this flexibility, the scattering matrix in the new base pair defined by U_{2A} and U_{2B} is related to that in xy and x'y' by a unitary similarity transformation as shown in equation (1.141):

$$[U_{2A}]^{-1} = [U_{2A}]^{*T} \Rightarrow \underline{E}'_{s} = [U_{2B}] \cdot [S] \cdot [U_{2A}]^{*T} \underline{E}'_{i}$$
(1.141)

The form of this transformation suggests that a singular value decomposition (SVD) of the matrix [S] may be used to simplify the transformation caused by [S] itself. According to the SVD (see Appendix 1), any complex matrix [S] may be written in terms of two unitary matrices [U] and [V] and a diagonal matrix

 $[\Gamma]$ as shown in equation (1.142):

$$[S] = [U] \cdot [\Gamma] \cdot [V]^{*T} = [U] \cdot \begin{bmatrix} \gamma_1 & 0 \\ 0 & \gamma_2 \end{bmatrix} \cdot [V]^{*T} \begin{cases} [U] \cdot [U]^{*T} = [I] \\ [V] \cdot [V]^{*T} = [I] \end{cases}$$
(1.142)

In the context of the scattering matrix, this result means that there always exists a pair of orthogonal bases for which the scattering matrix is diagonal; that is, has zero cross-coupling between states. We shall see that these represent the condition for maximum scattering cross-section of the object, and that for the special case of backscatter the matrices [U] and [V] are simply related so that a *single* antenna can be used to obtain this maximum backscatter signal. To see this, and to establish an important connection between the SVD and optimization based on Lagrange multipliers, consider solving the problem of maximising the received scattered power. To do this we first establish a Lagrangian function L, as shown in equation (1.143):

$$L = \underline{E}_{2}^{*T} [S] \underline{E}_{1} + \lambda_{1} (\underline{E}_{1}^{*T} \underline{E}_{1} - 1) + \lambda_{2} (\underline{E}_{2}^{*T} \underline{E}_{2} - 1)$$
(1.143)

Here we see in the first term the received complex voltage when we transmit a state \underline{E}_1 and receive with an antenna with polarisation \underline{E}_2 . In addition we have two constraint equations (and two corresponding unknown Lagrange multipliers) to ensure that the antenna polarisation vectors \underline{E} are unitary. To solve for the maximum received power LL^* we can sidestep the need to expand the product directly and instead set the partial derivatives of L and L^{*T} separately to zero, as shown in equation (1.144). (We shall use this same strategy when solving coherence optimization in polarimetric interferometry in Section 6.2.)

$$\frac{\partial L}{\partial \underline{E}_{2}^{*T}} = [S] \underline{E}_{1} + \lambda_{2} \underline{E}_{2} = 0$$

$$\frac{\partial L^{*T}}{\partial \underline{E}_{1}^{*T}} = [S]^{*T} \underline{E}_{2} + \lambda_{1}^{*} \underline{E}_{1} = 0$$

$$\Rightarrow [S]^{*T} [S] \underline{E}_{1} = \lambda_{1}^{*} \lambda_{2} \underline{E}_{1}$$
(1.144)

Hence we see in the optimum case that \underline{E}_1 is an eigenvector of the Hermitian matrix $[S]^{*T}$ [S], while \underline{E}_2 is an eigenvector of the matrix $[S][S]^{*T}$. In addition, as these matrices are Hermitian they each have a pair of orthogonal eigenvectors. These are just the left and right singular vectors of the [S] matrix, as easily verified from equation (1.142). In the context of radar polarimetry, the Hermitian matrix formed from products such as $[S]^{*T}[S]$ is called the Graves power matrix (Graves, 1956), as it contains information through its eigenvalues and eigenvectors on the variation of scattered power by the object. Note that only if [S] is symmetric are the matrices $[S]^{*T}$ [S] and $[S][S]^{*T}$ equal and the left and right singular vectors equal. In this case the maximum scattered power is formed from a copolarised signal (the same antenna for transmit and receive). We shall return to this issue in the context of backscatter, but first turn to consider some elementary symmetry properties of the general scattering amplitude matrix.

One basic issue of importance in polarimetry is the identification of any transformations of the problem in Figure 1.14, such that if we know the [S] matrix for the original system then can we predict its form for the new problem without the need for recalculation. This operation would then constitute a

symmetry operation of the system. We have already seen one such example: the problem of changing the incident wave polarisation. We have seen that we can easily predict the polarisation of the scattered wave using C2 symmetry coupled to matrix multiplication. However, this does not exhaust all possibilities. There are several such symmetry transformations (van der Hulst, 1982; Cloude, 1995b). As a trivial example we can translate the transmitter A or receiver B along the vectors $\underline{\beta}_i$ and $\underline{\beta}_s$ and effect only a scalar phase change in $[S] \rightarrow e^{i\varphi}[S]$, where ϕ is related to the change of range between A and B. This forms the basis for a discussion of interferometry with polarised waves, to be discussed at length in Chapter 6. Other interesting transformations are offered by considering rotations and reflections of the original problem in the bisectrix and scattering planes as we now consider.

We start by considering the matrix **S** (dispensing with square-bracket notation for the moment, for convenience) to be the 'mother' configuration, with complex elements a, b, c, and d. There are then three important 'daughter' matrices we can write for completely different scattering problems, as shown in equation (1.145):

$$S = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \Rightarrow \begin{cases} S_{\alpha} = \begin{bmatrix} a & -c \\ -b & d \end{bmatrix} \\ S_{\beta} = \begin{bmatrix} a & -b \\ -c & d \end{bmatrix} \quad a, b, c, d \in C \quad (1.145) \\ S_{\gamma} = \begin{bmatrix} a & c \\ b & d \end{bmatrix}$$

These matrices arise from the following fundamental geometrical transformations (see Figure 1.14):

- S_{α} Rotate the original problem by π about the bisectrix.
- S_{β} Mirror the original problem with respect to the scattering plane.
- S_{ν} Mirror the original problem with respect to the bisectrix plane.

The first of these, S_{α} , follows from the vector reciprocity theorem for electromagnetic waves. This theorem relates the exchange of transmitter and receiver positions, and states that if we transmit a polarisation state \underline{P}_A from A, then the component polarised in the \underline{P}_B direction at B is equal to the \underline{P}_A component of the scattered radiation when we illuminate the same object from B with polarisation \underline{P}_B . Figure 1.15 shows the reciprocal problem to that in Figure 1.14.



Fig. 1.15 Definition of the reciprocal scattering problem

Note how the incident and scattered wave vectors have changed direction. The scattering matrices for the two cases are then related by the reciprocity theorem as shown in equation (1.146):

$$\underline{\underline{E}}_{S}^{A} = \begin{bmatrix} S\left(\underline{\beta}_{i} \quad \underline{\beta}_{s}\right) \end{bmatrix} \cdot \underline{\underline{P}}_{A} \\
\underline{\underline{E}}_{S}^{B} = \begin{bmatrix} S\left(-\underline{\beta}_{i} \quad -\underline{\beta}_{s}\right) \end{bmatrix} \cdot \underline{\underline{P}}_{B} \\
\frac{\underline{P}_{B}}{\text{Reciprocity Theorem}} \\
\Rightarrow \begin{bmatrix} S\left(\underline{\beta}_{i} \quad \underline{\beta}_{s}\right) \end{bmatrix} = \begin{bmatrix} S\left(-\underline{\beta}_{i} \quad -\underline{\beta}_{s}\right) \end{bmatrix}^{T}$$
(1.146)

Note that the y and y' coordinates have been reversed in direction for the problem in Figure 1.15 compared to the master problem in Figure 1.14. To account for this we need to invert the y axis so that equation (1.146) becomes

$$\begin{bmatrix} S\left(\underline{\beta}_{i} \quad \underline{\beta}_{s}\right) \end{bmatrix} = \begin{bmatrix} 1 & 0\\ 0 & -1 \end{bmatrix} \begin{bmatrix} S\left(-\underline{\beta}_{i} & -\underline{\beta}_{s}\right) \end{bmatrix}^{T} \begin{bmatrix} 1 & 0\\ 0 & -1 \end{bmatrix}$$
$$\Rightarrow S = \begin{bmatrix} a & b\\ c & d \end{bmatrix} \Rightarrow S_{\alpha} = \begin{bmatrix} a & -c\\ -b & d \end{bmatrix} \qquad (1.147)$$

This leads directly to S_{α} in equation (1.145). The other two daughter problems are more easily derived: S_{β} follows from an inversion of one coordinate, while S_{γ} follows from successive application of α and β .

The reciprocity relation is especially important for backscatter problems, when the transmitter and receiver are at the same position. In this case it follows that $S = S_{\alpha}$, and from equation (1.147) this is only possible if the original 'mother' scattering matrix has the following form:

$$[S]_{backscatter} = \begin{bmatrix} a & b \\ -b & d \end{bmatrix}$$
(1.148)

This is called the backscatter theorem, and demonstrates that in this special case the S matrix has only three independent elements. Note that while there exists a class of non-reciprocal backscatter problems, these are not common in the remote sensing of natural land and sea surfaces (an important exception being propagation through the ionosphere when Faraday rotation occurs; see Section 1.2.2). The reciprocity theorem is therefore an important basic symmetry, widely employed in radar backscatter systems for information extraction and calibration.

It is of interest to now reconsider an SVD of the backscatter matrix of equation (1.148) and to relate the problem to that of matched antenna illumination. To do this we must first consider a correction to account for the differences between transmitter and receiver antenna coordinates—a generalization of that following from equation (1.147).

1.4.2 Back and forward scattering alignment (BSA and FSA) systems

In the previous section we considered a description of the scattering matrix in a coordinate system that in a sense 'follows' the wave propagation direction. This is called the forward scattering alignment or FSA system. In radar and antenna

 $\underline{\mathbf{h}} = \mathbf{h}_{\theta} \underline{\theta} + \mathbf{h}_{\omega} \underline{\phi}$ complex effective length of the antenna

studies, however, a different system is employed. This is called the backscatter alignment or BSA system. In this section we motivate the need for such a change and outline the main differences between FSA and BSA descriptions of scattering (Kostinski, 1986).

When an antenna receives a signal it converts the incident electric field (with units of Volts/m) into a circuit measurement in Volts (Collin, 1985; Mott, 1992). The transfer function from field to circuit therefore has the units of length (m). Given the C2 freedom of propagating waves, this transfer function must be a two-component complex vector \underline{h} , called the complex effective length of the antenna. This basic antenna relationship is defined in Figure 1.16. If the antenna and field are orthogonally polarised then zero voltage will be received, regardless of the amplitude of the incident field. At the opposite extreme, when field and antenna and field. In this section we consider the details of this coupling process and how it relates to the structure of the scattering matrix.

An obvious starting point, common to vectors in C2, is to assume that the open circuit voltage is given by the Hermitian product between receiving antenna complex effective length \underline{h} and the incident polarisation vector \underline{E}_i . This has the desired property of being zero for orthogonality and a real maximum value under the condition $\underline{h} = \underline{E}_i$

The first problem we face is that in active systems antennas are used both for transmission and reception, and we must therefore take care over the definition of antenna polarisation vector \underline{h} . In common with engineering standards we define the antenna polarisation in *transmit* mode. Hence a +45-degree linearly polarised antenna has a normalized effective length $\underline{h} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix}$, defined in a right-hand coordinate system as shown on the top line of Figure 1.17. Now consider an experiment where two *identical* such antennas are used as the basis for a communication system. Figure 1.17 shows a schematic of such an arrangement, and here we see that, although identical, when used in transmit/receive mode the antennas are orthogonal and the communication channel is null. Clearly, in this case we cannot assume the matched condition as $\underline{E}_i = \underline{h}$.

The source of this problem can be traced to a mismatch in right-hand coordinate systems for transmit and receive. We see that one component of the *receiver* is 180-degree phase-shifted compared to that of the *transmitter*, due to the z-coordinate reversal inherent in point-to-point communication. To compensate for this we must introduce a minus sign or differential 180-degree phase shift between the polarisation channels, as shown in Figure 1.17. With this in place the Hermitian product between effective length and incident field produces the correct null result. Since this phase correction is in the form of a matrix product it can be incorporated into the [S] matrix definition. First, however, we must also consider the fact that the waves propagate in different directions, from which a further complication arises when dealing with elliptical polarisations. Fig. 1.16 Complex effective length of an antenna



Fig. 1.17 Transmission and reception between $+45^{\circ}$ polarised antennas (top) and left-hand circularly polarised antennas (bottom), showing the need for coordinate reversal and conjugation in point-to-point communication between antennas

To illustrate the nature of the problem we again consider a communication link, established this time between two identical left circularly polarised antennas. The lower part of Figure 1.17 shows the details of this arrangement. In this case the 180-degree coordinate phase correction with Hermitian product again leads to a predicted mismatch between antennas. However, this is incorrect. Identical circularly polarised antennas remain matched when used in point-to-point communication. The reason for this change of behaviour when switching from linear to circular is that while the sense of circular polarisation does change with the coordinate inversion, it *also* changes with time reversal. This sense reversal is mathematically represented by the complex conjugate operator applied to the incident field, since if <u>P</u> is a polarisation vector then <u>P</u>* has the same polarisation ellipse, but with opposite sense.

With these coordinate transformations in place, the voltage received by an antenna with effective length \underline{h}_2 when the scatterer is illuminated by a wave originating from an antenna with complex length \underline{h}_1 , when both \underline{h}_1 and \underline{h}_2 are defined in the *transmitting* mode of the antennas, can be written as shown in equation (1.149):

$$V_{oc} = \left(\underline{h}_{2}^{*T}\right)^{*} \cdot \begin{bmatrix} 1 & 0\\ 0 & -1 \end{bmatrix} \cdot \begin{bmatrix} a & b\\ c & d \end{bmatrix} \cdot \underline{h}_{1} = \underline{h}_{2}^{T} \left[S\right]_{sensor} \underline{h}_{1}$$
(1.149)

Consequently, it is common practice in radar—where backscatter problems predominate and so the same antenna is used for transmission and reception—to write the open circuit voltage in the backscatter alignment or BSA system as a complex inner rather than Hermitian product of a coordinate reflected scattering

matrix, as shown in equation (1.150).

$$V_{oc} = \underline{h}^{I} [S]_{sensor} \underline{.h}$$

$$\Rightarrow [S]_{BSA} = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} [S]_{FSA} = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \underline{.} \begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} a & b \\ -c & -d \end{bmatrix}$$
(1.150)

Also shown in equation (1.150) is the relationship between the scattering amplitude matrix in the sensor coordinates (which becomes the backscatter alignment system or BSA system in backscatter) and the wave or FSA coordinates. Note the phase change of elements of [S]. Care is therefore required in always specifying which coordinate system is being employed in a description of scattering. In the engineering literature the sensor or BSA system is almost universally used, while in optics and physics the conventional FSA or wave system is preferred.

In equation (1.148) we saw that from the backscatter theorem, b = -c in backscatter. It follows from equation (1.150) that the BSA matrix becomes complex symmetric; that is, b = c, and the received backscatter voltage for reciprocal problems can be written in the most general case, as shown in equation (1.151):

$$V_{oc} = \underline{h}_{2}^{T} [S]_{BSA} \underline{h}_{1} = \underline{h}_{2}^{T} \begin{bmatrix} a & b \\ b & -d \end{bmatrix} \underline{h}_{1}$$
(1.151)

This is the most common form of this equation used in radar backscattering theory. Some examples will be used to illustrate its application.

1.4.2.1 Example 1 : specular backscatter matrix

Consider the simple problem of normal incidence reflection from a flat surface, as shown schematically in Figure 1.18. The reflected field is given by the Fresnel equations at normal incidence, which yield zero crosspolarisation and a 180-degree phase difference between copolar coefficients (see Section 3.1.1). Hence we can write the reflection matrix $[R_{ef}]$ in the FSA wave coordinate systems as shown in equation (1.152):

$$\begin{bmatrix} R_{ef} \end{bmatrix} = A \begin{bmatrix} 1 & 0\\ 0 & -1 \end{bmatrix}$$
(1.152)

where A is a constant that depends on the dimensions of the surface relative to a wavelength and on the dielectric constant of the surface material. Note that the matrix $[R_{ef}]$ turns the reflection problem into an equivalent point-topoint communication problem, and so the received voltage for transmission and reception from an antenna with effective length <u>h</u> can be written as shown in equation (1.153):

$$V_{oc} = \underline{h}^{T} \cdot \underline{\underline{E}}_{i} = \underline{h}^{T} \cdot \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \underline{\underline{E}}_{s} = A\underline{\underline{h}}^{T} \cdot \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \underline{\underline{E}}_{t} = A\underline{\underline{h}}^{T} \cdot \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \underline{\underline{h}}$$
(1.153)

Hence the scattering matrix in the sensor or BSA coordinates is a multiple of the identity matrix, and we have the following important examples of antenna



Fig. 1.18 Specular backscatter from a surface at normal incidence

match and mismatch: namely, that circular polarisation yields zero received voltage, while ± 45 -degree linear is matched for maximum signal strength (as incidentally are all linear incident polarisations).

$$[S]_{BSA} = A \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \Rightarrow \begin{cases} \underline{h} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ \pm i \end{bmatrix} \Rightarrow V_{oc} = 0 \\ \underline{h} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ \pm 1 \end{bmatrix} \Rightarrow V_{oc} = A \end{cases}$$
(1.154)

1.4.2.2 Example 2: dihedral scattering matrix

A second important way in which specular backscattering can occur is in dihedral retroreflection, as shown schematically in Figure 1.19. In this case we have a 90-degree corner formed from two perfect conductors (PC), and the ray paths indicate a strong specular return to the transmitter through two 45-degree reflections at each surface. From a polarisation point of view the reflection matrix is formed from the product of the reflection matrices at A and B. In general, the Fresnel coefficients change with angle of incidence onto the surface (see Section 3.1.1), and away from normal incidence are not given by the simple form of $[R_{ef}]$ in equation (1.152). We shall consider the more general case later, but here concentrate on perfect conductors only. For this special case, $[R_{ef}]$ applies for all angles, and using a cascade of matrices of the form of equation (1.152): we can generate the normalized reflection matrix shown in equation (1.155):

$$\begin{bmatrix} R_{ef} \end{bmatrix} = \begin{bmatrix} R_{ef} \end{bmatrix}_A \begin{bmatrix} R_{ef} \end{bmatrix}_B = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$
(1.155)

The voltage received for an antenna with effective length \underline{h} in the BSA system is then of the following form:

$$V_{oc} = \underline{h}^T \cdot \underline{E}_i = \underline{h}^T \cdot \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \underline{E}_t = \underline{h}^T \cdot \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \cdot \underline{h}$$
(1.156)

and so the scattering matrix for a dihedral has the form shown in equation (1.157):

$$[S]_{BSA} = A \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \Rightarrow \begin{cases} \underline{h} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ \pm i \end{bmatrix} \Rightarrow V_{oc} = A \\ \underline{h} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ \pm 1 \end{bmatrix} \Rightarrow V_{oc} = 0 \end{cases}$$
(1.157)

We now see that a circularly polarised antenna is matched to the return signal, while a 45-degree linearly polarised antenna is mismatched, rather like the communication system of Figure 1.17. Clearly the pattern is now set for higherorder scattering processes. The BSA scattering matrix for order N scattering from a set of PC surfaces is denoted as follows:

$$[S]_{BSA} = \begin{bmatrix} 1 & 0\\ 0 & -1 \end{bmatrix}^{2N+1}$$
(1.158)



Fig. 1.19 Dihedral retro-reflection



Fig. 1.20 Image of a trihedral corner reflector used for polarimetric calibration

so for *N* odd (1,3,5 and so on) there is zero phase difference between copolar elements, while for *N* even (2,4,6 and so on) there is π phase difference. The generalization of these ideas to dielectric surfaces is dealt with in Section 3.1, and leads us eventually to use of the scattering alpha parameter to accommodate the changes in boundary conditions for general surfaces.

As an important example of equation (1.158), consider the case N = 3, which occurs for a trihedral corner reflector, widely used for radar backscatter calibration, as it combines high radar cross-section with a broad radiation pattern, so simplifying alignment with the sensor. (Contrast this with a flat plate N = 1, which can always be made bright by increasing its size but only at expense of the radiation pattern, which becomes more localized around the specular direction in a narrow pencil beam.) Figure 1.20 shows an image of a metallic trihedral N = 3 reflector. From a polarisation point of view, this scatterer behaves the same as specular normal surface reflection (N = 1), and has a BSA scattering matrix equal to the 2×2 identity matrix. Hence it too is 'blind' to circularly polarised antennas, regardless of its physical size.

1.4.3 Singular value analysis of the scattering matrix

In the previous section we saw that the boundary conditions at perfectly conducting surfaces lead to a set of 'blind polarisations' for even and odd number of reflections (circular for odd and 45-degree linear for even). In order to extend these ideas to arbitrary materials and shapes we now turn to consider a generalization of blind polarisations through a singular value decomposition (SVD) of the backscattering matrix in the sensor or BSA coordinate system. To do this we note that when we change the polarisation basis of the incident field by a unitary matrix U, then the open circuit voltage can be written as shown in equation (1.159), from which we see that the backscattering matrix transforms not as a similarity transformation but as a unitary congruent transformation (involving only the transpose and not the transpose conjugate) (Luneburg, 1996).

$$V_{oc} = \underline{h}^{T} [U]^{T} [S]_{sensor} [U] \underline{h} \Rightarrow [S]' = [U]^{T} [S]_{sensor} [U]$$
(1.159)

On the other hand, for complex symmetric matrices the two unitary components [U] and [V] of the SVD are related, since we have the following identity:

$$[S] = [U] \cdot [\Gamma] \cdot [V]^{*T} = [S]^{T} = [V]^{*} \cdot [\Gamma] \cdot [U]^{T}$$

$$\Rightarrow [U] = [V]^{*}$$

$$\Rightarrow [S] = [U] \cdot [\Gamma] \cdot [U]^{T}$$
(1.160)

and so we see from the SVD that a symmetric [S] matrix can always be diagonalized by a congruent unitary transformation, just as we found for the scattering matrix in the sensor coordinates (equation (1.159)). Some examples will help illustrate how this formalism can be used in practice. Firstly we consider the simple case of rotation of the reference plane of polarisation for backscattering by a horizontal dipole and a dihedral reflector. The scattering amplitude matrices for these two cases can be derived from standard similarity transformations involving a plane rotation matrix, as shown in equation (1.161):

$$[S_{\theta}] = \begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \cdot \begin{bmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{bmatrix} = \begin{bmatrix} \cos^{2} \theta & \cos \theta \sin \theta \\ \cos \theta \sin \theta & \sin^{2} \theta \end{bmatrix}$$
$$[S_{\theta}] = \begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \cdot \begin{bmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{bmatrix} = \begin{bmatrix} \cos^{2} \theta - \sin^{2} \theta & 2\cos \theta \sin \theta \\ 2\cos \theta \sin \theta & \sin^{2} \theta - \cos^{2} \theta \end{bmatrix}$$
(1.161)

Some care is required, however, when considering change to elliptical polarisation bases. As an important special case we consider the form of the scattering matrix in the circular polarisation base, obtained by a *congruent* unitary transformation of the linear [S] matrix, as shown in equation (1.162). Note how the left and right side transformation matrices are now identical. The reason for this is the coordinate manipulations inherent in the BSA system.

$$[S]_{circ} = \begin{bmatrix} S_{LL} & S_{LR} \\ S_{RL} & S_{RR} \end{bmatrix} = \frac{1}{2} \begin{bmatrix} 1 & i \\ i & 1 \end{bmatrix} \begin{bmatrix} S_{HH} & S_{HV} \\ S_{HV} & S_{VV} \end{bmatrix} \begin{bmatrix} 1 & i \\ i & 1 \end{bmatrix}$$
$$= \frac{1}{2} \begin{bmatrix} S_{HH} - S_{VV} + 2iS_{HV} & i(S_{HH} + S_{VV}) \\ i(S_{HH} + S_{VV}) & S_{VV} - S_{HH} + 2iS_{HV} \end{bmatrix}$$
(1.162)

For example, consider again the rotated dihedral matrix now expressed in the circular base (substitute the lower example in equation (1.161) into (1.162)). The result is shown in equation (1.163). Now we see a much-simplified diagonal form with rotation influencing only the phase between LL and RR polarisations. This is an example of diagonalization via the SVD.

$$[S]_{circ} = \begin{bmatrix} \cos^2 \theta - \sin^2 \theta + i2 \sin \theta \cos \theta & 0\\ 0 & \sin^2 \theta - \cos^2 \theta + i2 \sin \theta \cos \theta \end{bmatrix}$$
$$= \begin{bmatrix} \cos 2\theta + i \sin 2\theta & 0\\ 0 & -\cos 2\theta + i \sin 2\theta \end{bmatrix} = \begin{bmatrix} e^{i2\theta} & 0\\ 0 & -e^{-i2\theta} \end{bmatrix}$$
(1.163)



Fig. 1.21 Schematic representation of the XPOL and COPOL nulls

The main conclusion is that for any given symmetric complex scattering matrix we can always find an orthogonal base 'ab' such that [*S*] is diagonal in that base; that is, there is zero crosspolarisation. Incidentally, this is reminiscent of eigenpolarisations in wave propagation (Section 1.2.2), except here we are dealing with the physically distinct case of wave scattering. These states are sometimes confusingly called eigenstates, but should more accurately be called crosspolar nulls, or often abbreviated to XPOL nulls of the scattering matrix, as originated by Kennaugh (1952) and further developed by Huynen (1987) and Boerner (1981). Note that for symmetric matrices there are always two orthogonal XPOL nulls. They have two very interesting properties, as follows:

- The XPOL nulls maximize the backscattered signal (for a fixed antenna polarisation), and thus maximize the detectability of the object in the presence of noise.
- The XPOL nulls remain unchanged on scattering from the object, and thus 'carry' information about any symmetry properties of the object.

Figure 1.21 shows a schematic representation of this invariance property of the XPOL nulls. By definition, when P is an XPOL null then the scattered field is also P polarised (in the BSA coordinates) as shown. For example, when the scatterer has an axis of symmetry in the plane of polarisation then the XPOL nulls are linear polarisations aligned and perpendicular to this symmetry axis. In this way we can establish the orientation of an unknown symmetry axis by using an SVD of the measured scattering matrix. But how does this diagonalization process link with the idea of blind polarisations? To see this, we now seek incident polarisation states that are not invariant but rather transformed into their orthogonal state on scattering. These states would then by definition constitute the blind polarisations of the system. In distinction to the diagonalization process, these states are called copolar nulls or simply COPOL nulls, and we shall now show that there are always two such states for all [S] matrices, although unlike XPOL nulls they are only orthogonal under special circumstances. The trick is to solve for COPOL nulls starting from the diagonal form of the [S] matrix, as follows. Recall that the XPOL nulls are always orthogonal (forming a base we shall call 'ab') and so we can always write the scattering matrix in this base in diagonal form, as shown in equation (1.164):

$$[S] = \begin{bmatrix} \gamma_a & 0\\ 0 & \gamma_b \end{bmatrix}$$
(1.164)

where γ_a and γ_b are the complex singular values of [S]. We can represent any polarisation state as a linear combination of these XPOL nulls, that is, $\underline{E} = E_a \underline{a} + E_b \underline{b}$, and hence for COPOL nulls we seek solutions of the form shown in equation (1.165):

$$\begin{bmatrix} E_a & E_b \end{bmatrix} \cdot \begin{bmatrix} \gamma_a & 0 \\ 0 & \gamma_b \end{bmatrix} \cdot \begin{bmatrix} E_a \\ E_b \end{bmatrix} = 0$$
$$\Rightarrow E_a^2 \gamma_a + E_b^2 \gamma_b = 0$$
$$\Rightarrow \frac{E_a}{E_b} = \pm \sqrt{-\frac{\gamma_b}{\gamma_a}}$$
(1.165)

This shows that there are two COPOL nulls (with a complex polarisation ratio given by the plus and minus square root), and that generally they are *not* orthogonal to each other. As simple examples we return to the matrices for specular and dihedral scattering (equations (1.154) and (1.157)). These are both diagonal in the given coordinate system, and their singular values are therefore easily obtained by inspection. The results for the COPOL nulls are then given as shown in equation (1.166), which we see agrees with our test cases considered in equations (1.154) and (1.157). Equation (1.165) then generalizes this result to arbitrary scattering matrices.

$$[S] = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \xrightarrow{\text{copol nulls}} \frac{E_x}{E_y} = \pm i$$

$$[S] = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \xrightarrow{\text{copol nulls}} \frac{E_x}{E_y} = \pm 1$$
 (1.166)

Finally, we note that there are three important invariants of the scattering amplitude matrix under unitary congruent transformations, as shown in equation (1.167):

$$[S] = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \xrightarrow{\text{invariants under } U^{T}SU} \begin{cases} \det([S]) = ad - bc & I \\ Span([S]) = |a|^{2} + |b|^{2} + |c|^{2} + |d|^{2} & II \\ i(c-b) & III \end{cases}$$

$$(1.167)$$

The first follows directly from the use of special unitary matrices in establishing the change of basis matrix. The second allows us to define the norm or magnitude of a scattering matrix, and will be used to define scattering amplitude vectors. The third assures us that if the scattering matrix is symmetric in one base, it remains so in all bases. To prove this, consider an expansion of the amplitude matrix in terms of the complete and orthogonal Pauli matrix set, as shown in equation (1.168), where the square root of two normalization is included to ensure that $|\underline{k}|^2 = Span([S])$ from invariant II.

$$[S] = (k_o \sigma_0 + k_1 \sigma_1 + k_2 \sigma_2 + k_3 \sigma_3) \quad \Rightarrow [S] = \sum_{i=0}^3 k_i \sigma_i \qquad k_i = \frac{1}{\sqrt{2}} Trace([S] \sigma_i)$$
(1.168)

It follows from properties of the Pauli matrices that we can form the identity

$$[U]^{T}[S][U] = [U]^{T}(k_{o}\sigma_{0} + k_{1}\sigma_{1} + k_{2}\sigma_{2} + k_{3}\sigma_{3})[U] \implies [U]^{T}\sigma_{3}[U] = \sigma_{3}$$
(1.169)

which leads to invariant III shown in equation (1.167).



Fig. 1.22 Three-stage decomposition of combined wave propagation and scattering

1.4.4 Combined wave propagation and scattering effects

Now we can combine the Jones propagation calculus of Section 1.2.4 with the above scattering matrix results to obtain a general expression for the scattering matrix of an object observed in the presence of non-trivial wave propagation effects. This is important in many practical applications, such as radio wave propagation through the ionosphere for space-based radar Earth observation (Bickel, 1965) and scattering from objects embedded in chiral materials such as occur in light scattering by biological systems (Ablitt, 1999, 2000).

To develop this we consider a three-stage process as shown schematically in Figure 1.22. The transmitter is located at A, and the first stage involves propagation of the wave from A to the object at B. This is represented by a 2×2 complex propagation matrix $[M_z(A,B)]$ which, as we saw in Section 1.2, needs be neither unitary nor homogeneous:

Stage I :
$$\underline{E}^B = [M_z(A, B)] \underline{E}^A$$

The second stage involves scattering of the wave by the target. This can be represented by a 2×2 complex scattering matrix [S], which we express in the sensor or BSA coordinate system as follows:

Stage II :
$$\underline{E}_{S}^{B} = [S_{BSA}] \underline{E}^{B} = [S_{BSA}] \cdot [M_{z}(A, B)] \underline{E}^{A}$$

The third stage involves propagation back through the medium from B to A. As we are remaining in the A sensor coordinate system then the reciprocity theorem implies that $[M_z(A,B)] = [M_z(B,A)]^T$ (equation (1.146)), so that we can write the following:

Stage III :
$$\underline{E}_{s}^{A} = [M_{z}(B,A)] \cdot \underline{E}_{s}^{B} = [M_{z}(A,B)]^{T} \cdot [S_{BSA}] \cdot [M_{z}(A,B)] \underline{E}^{A}$$

Consequently we can write the general combination of propagation and scattering in the BSA system as shown in equation (1.170):

$$[S]_{observed} = [M_z]^T [S_{BSA}] [M_z]$$
(1.170)

This is our desired result. It shows that the observed scattering matrix for a combined reciprocal propagation and scattering channel is, in the sensor BSA coordinate system, again obtained as a congruent transformation of the target scattering matrix by the propagation channel matrix. This represents a generalization of the change of basis, which resulted in a unitary congruent transformation.

Having now obtained these basic algebraic results for the complex amplitude matrix, its form under change of base and its fusion with propagation, we now turn to consider a geometrical interpretation of these transformation properties.
1.5 Geometry of the scattering matrix

In Section 1.3 we showed how all states of wave polarisation can be mapped onto the surface of the Poincaré sphere. We saw that there are three possible coordinate systems used to represent polarisation state on this sphere—two angle-based and one Cartesian—leading to the Stokes four-vector representation. Here we consider the implications of these results for a geometrical interpretation of the scattering amplitude matrix [*S*].

In general terms the scattering matrix represents a polarisation state transformer. Each possible input state, represented by a point on the sphere, is mapped into a second point with contracted radius (assuming the target is passive and has no inherent signal gain). The details of this mapping are of interest as a stage towards information extraction from the scattering matrix. For example, we shall find that some points are fixed under this transformation and that their relative location on the sphere is an indicator of target symmetry.

1.5.1 Polarisation signatures

The first stage of development is to be able to map variations in scattered amplitude with movement of points over the Poincaré sphere. This can be achieved in several ways, all of which require map projections of the spherical coordinates onto a plane. Three important such methods are summarized in Figure 1.23 (van Zyl, 1987; Woodhouse, 2003). On the far left we show the polarimetric signature method, in the centre a polar projection of the sphere, and on the right a hybrid polar system that combines the advantages of the other two into a convenient form for visualization.

One of the first such methods developed was simply to map the surface into a bounded rectangular region, as shown on the left in Figure 1.23. Each point inside the rectangle is then mapped into a transmit/receive antenna configuration and the resulting scattered power obtained from the voltage formula of equation (1.151). A common configuration is to choose a copolar representation, where the same point on the Poincaré sphere represents the transmit *and* receive antenna polarisation states (in the BSA convention). A second popular choice is to map the crosspolar channel variation. Here a point on the sphere is chosen for transmit and its antipodal partner for reception. There are two main problems with this rectangular representation. The first is the artificial discontinuity introduced in the θ coordinate by using a rectangular region. The second is the spreading of the circular polarisation poles of the sphere into



Fig. 1.23 Planar mapping representations of the surface of the Poincaré sphere

extended linear regions. This distorts the information for points located away from the equator. One advantage, however, is the clear separation of left and right sense polarisations into upper and lower halves of the diagram. This permits easy visualization of asymmetry due to calibration errors or scattering from asymmetric objects such as helices.

To overcome the distortion limitation, use is often made of a polar projection of the sphere, as shown in the centre of Figure 1.23. Here the pole is located at the centre of the image, and the polar θ coordinate is represented in a more natural way, without the need for any artificial discontinuities. However, one remaining problem is that the double angle representation of the Poincaré sphere leads to upper and lower hemispheres being mapped on top of each other. Hence there is no discrimination in this diagram between left and right sense polarisations. One compromise solution is to employ the representation shown on the right-hand side of Figure 1.23. Here we maintain the polar representation with circular polarisation at the centre, but now map θ in the range $-\pi/2$ to $\pi/2$. This frees the second half of the polar plot for the opposite sense of polarisation. In this way we can map both senses and all linear states onto a simple polar diagram (Woodhouse, 2003).

Figure 1.24 shows examples of all three representations for a set of sample matrices. The first three are the backscattering matrices corresponding to



Fig. 1.24 Copolarised backscatter power signatures for canonical scattering matrices

the Pauli set. The first of these is just the identity matrix, and we can see the characteristic copolar nulls at left and right circular polarisations. Maximum response is obtained for all linear polarisations, and there is no left/right dependency of scattered power. This represents the most symmetric of scattering systems, showing dependency on neither orientation nor sense of polarisation. The second and third Pauli matrices demonstrate rotation dependence. Here the copolar nulls are located at 45-degree and 0-degree linear polarisations respectively. Maximum backscatter is now achieved for a set of elliptical polarisations lying along a great circle of the sphere, passing through the poles corresponding to left and right circular polarisation. Again, however, we note that there is no left/right dependence on backscatter. These clearly show how the third Pauli matrix can be obtained from the second by coordinate rotation through 45 degrees. The next scattering matrix shown is that for a helix. This is chosen to illustrate an object that scatters one sense of circular polarisation in preference to the other. The asymmetry in the copolarised plot is clearly seen in the first and third diagrams. Finally we show the scattering matrix for a dipole oriented at 45 degrees to the horizontal. The orientation of the scatterer is clearly seen directly in the third representation. These final two examples demonstrate the benefit of the hybrid representation for visualising both rotation and sense preference in scattering.

Of particular interest are the extreme values of these scattering functions. These can be analytically evaluated with the help of a singular value decomposition of the [S] matrix. We showed in Section 1.4.3 that for an arbitrary scattering matrix there exist two pairs of orthogonal states for which the scattered power is maximized (XPOL nulls). We further showed that in backscatter these maxima always correspond to copolar combinations and will therefore appear in the polarisation charts of Figure 1.24. Also, there are two (non-orthogonal) zero minimum points which for backscatter also correspond to zeros in the copolar function. In backscatter, therefore, copolar and crosspolar plots are useful for assessing the maximum variation of scattering with changes in polarisation. However, in the general non-symmetric scattering matrix case, as arises in bistatic scattering, such functions are not guaranteed to contain global maxima and minima.

1.5.2 The polarisation fork

The extreme points in Figure 1.24 correspond directly to the XPOL and COPOL nulls. These four special points on the Poincaré sphere then form the basis for a geometrical interpretation of the general transformation of polarisation state, called the polarisation fork (Kennaugh, 1952; Huynen, 1987). The two XPOL nulls are maxima (minima) in the copolar (crosspolar) signature plots and, always being orthogonal, correspond geometrically to antipodal points lying at opposite ends of a diameter of the Poincaré sphere. These are shown as P and Q in Figure 1.25. We showed in equation (1.165) that the COPOL nulls are then obtained in the basis defined by PQ from the square root of the ratio of complex singular values of the scattering matrix. The square root having two solutions, there are then two of these, displaced symmetrically about the line PQ. These are shown as R and S in Figure 1.25. To see this we note that the unitary polarisation



Fig. 1.25 Polarisation fork geometry

vectors for R and S can be written as shown in equation (1.171):

$$P_{R,S} = \begin{bmatrix} \cos \alpha_w \\ \sin \alpha_w e^{i\delta} \end{bmatrix}_{PQ} \Rightarrow \tan \alpha_w e^{i\delta} = \pm \sqrt{-\frac{\gamma_P}{\gamma_Q}}$$
(1.171)

where γ are the complex singular values $|\gamma_P| \ge |\gamma_Q|$ and the α_w angle is defined as shown in Figure 1.25. Note that $\alpha_w \ge \frac{\pi}{2}$, and so very often the angle $2\gamma = \pi - 2\alpha_w$ is employed. γ is called the fork angle of the scattering matrix (Kennaugh, 1952; Huynen, 1970, 1987). We can then rewrite the diagonal scattering matrix as a function of the fork angle, as shown in equation (1.172):

$$[S] = \begin{bmatrix} \gamma_P & 0\\ 0 & \gamma_Q \end{bmatrix} = m e^{i\phi} \begin{bmatrix} \cot \gamma e^{i\frac{\chi}{2}} & 0\\ 0 & \tan \gamma e^{-i\frac{\chi}{2}} \end{bmatrix}$$
(1.172)

The angle between the singular vectors, χ , is called the skip angle. Note that the matrix on the right-hand side has unit determinant. This will provide an important link with the geometry of the Lorentz transformation in the next section.

We note from this analysis that the four points P, Q, R, and S all lie in a plane in the space of the Poincaré sphere. This structure is called the polarisation fork, and was first developed by Kennaugh (1952) and later studied by several authors, particularly Huynen (1970, 1987). Consequently, the angular parameters of this plane and fork geometry are often termed the Huynen parameters. Kennaugh also devised a simple set of geometrical rules for predicting the transformation of polarisation state by a scattering matrix. The rules employ the point I shown as the intersection of the chord joining the COPOL nulls and the diameter PO in Figure 1.25. For a given scattering matrix, the transformation of any incident polarisation can then be found by first inverting the point through I onto the sphere. This new point is then rotated by π about the diameter normal to PQ. The coordinates of the new point then correspond to the polarisation of the scattered wave. Applying these rules to the points P and Q themselves confirms that they remain unchanged on scattering, while R and S are both mapped into their antipodal orthogonal points as expected.

Two special classes of scattering can be defined based on the α angle. When the singular values are equal in amplitude, then $\alpha_w = 45$ degrees, the COPOL nulls are orthogonal, and I is located at the origin. Then, due to the symmetry, all polarisation states lying on a great circle formed by rotating the fork about the RS axis are also XPOL nulls. This is the case for the three Pauli matrices shown at the top of Figure 1.24, where the loci of maximum scattered power are seen as bands of white across the diagrams. This arises as the Pauli matrices have singular values that differ only in phase but are equal in amplitude.

The second important class occurs when the scattering matrix is singular (zero determinant) and consequently one of the singular values is zero. In this case $\alpha_w = 90$ degrees, R,I,S collapse to the point Q, and all polarisation states are transformed into the state P, regardless of their initial position on the sphere.

Such a device is a polariser, in that it scatters the same polarisation state for all input states. Important examples are shown in the lower two sections of Figure 1.24. The dipole and the helix both have singular scattering matrices, and in both cases the fork collapses to a line.

Note that a similar construct can be used for the case of non-symmetric scattering matrices. The complex singular values are still well defined, and thus the fork angle is also uniquely defined. The axis PQ now corresponds to the left singular vectors and is again a diameter of the sphere, as P and Q remain orthogonal. Hence the fork geometry is still maintained. However, now the scattered wave must be interpreted in a new P'Q' orthogonal base defined by the right singular vectors. Hence we have a fork for transmit and a second fork for reception, although the fork angle remains invariant.

1.5.3 Lorentz geometry and the scattering matrix

The Poincaré formulation is based on the geometry of the unit sphere, and thus ignores any changes in scattered amplitude. The copolar signatures of Figure 1.24 are generated essentially by using a form of Malus's law with a metric of $\cos^2 \phi/2$ (Hecht, 1997). Here ϕ is the angular separation of the initial and final polarisation states on the surface of the Poincaré sphere. However, this works only in a *relative* sense because we have defined reference points on the surface (the original point for copolar signatures plus the antipodal point for crosspolar). In other words, we must choose cross- or copolar projections for the signature. It would be better if we could devise a method of interpretation that did not require such a choice.

In more general terms we would like to devise a geometrical interpretation of the change in both amplitude and polarisation state. Such an approach can be developed based on the geometry of the Lorentz transformation. Just as there is an homomorphism between the group of special unitary matrices and that of proper rotations, there is also a homomorphic (2-to-1) mapping from the group of 2×2 complex unimodular matrices [Q] and pure Lorentz transformations (see Appendix 2 and Goldstein (1980)). In this section we explore the geometrical implications of this mapping for an interpretation of the scattering matrix.

We begin by again considering the diagonal matrix of singular values of [S]. We can rewrite this in unimodular form as shown in equation (1.173), where we also show explicitly the polar decomposition of [Q] into unitary and Hermitian matrices [U] and [H] respectively.

$$[S] = \begin{bmatrix} \gamma_P & 0\\ 0 & \gamma_Q \end{bmatrix} = a[Q] = a \cdot [U] \cdot [H]$$
$$= a \begin{bmatrix} e^{i\frac{\chi}{2}} & 0\\ 0 & e^{-i\frac{\chi}{2}} \end{bmatrix} \cdot \begin{bmatrix} \cot \gamma & 0\\ 0 & \tan \gamma \end{bmatrix}$$
$$= a \begin{bmatrix} e^{i\frac{\chi}{2}} & 0\\ 0 & e^{-i\frac{\chi}{2}} \end{bmatrix} \cdot \begin{bmatrix} e^b & 0\\ 0 & e^{-b} \end{bmatrix}$$
(1.173)

This is important, because it is the Hermitian component that leads directly to the new type of geometry. We saw in Section 1.3.3 that in order to include

amplitude information we need to extend the Pauli set by addition of the identity matrix and obtain a modified spin matrix. To see how matrices of the form [Q] change this spin matrix we form the triple matrix product shown in equation (1.174):

$$[R(\chi,\tau)] = \frac{1}{2}[Q] \cdot \begin{bmatrix} g_0 + g_1 & g_2 - ig_3 \\ g_2 + ig_3 & g_0 - g_1 \end{bmatrix} \cdot [Q]^{*T}$$
(1.174)

If we set the skip angle χ to zero and determinant to unity in equation (1.174), this transformation has the following simple canonical form:

$$\begin{bmatrix} g_0^{new} + g_1^{new} & g_2^{new} - ig_3^{new} \\ g_2^{new} + ig_3^{new} & g_0^{new} - g_1^{new} \end{bmatrix} = \frac{1}{2} \begin{bmatrix} e^b & 0 \\ 0 & e^{-b} \end{bmatrix} \cdot \begin{bmatrix} g_0 + g_1 & g_2 - ig_3 \\ g_2 + ig_3 & g_0 - g_1 \end{bmatrix} \cdot \begin{bmatrix} e^b & 0 \\ 0 & e^{-b} \end{bmatrix}$$
(1.175)

By direct expansion it follows that we can relate the Stokes vectors of the incident and scattered field by a 4×4 matrix [*M*], as shown in equation (1.176):

$$\begin{bmatrix} g_{0}^{new} \\ g_{1}^{new} \\ g_{2}^{new} \\ g_{3}^{new} \end{bmatrix} = \begin{bmatrix} m_{00} & m_{01} & m_{02} & m_{03} \\ m_{10} & m_{11} & m_{12} & m_{13} \\ m_{20} & m_{21} & m_{22} & m_{23} \\ m_{30} & m_{31} & m_{32} & m_{33} \end{bmatrix} \cdot \begin{bmatrix} g_{0} \\ g_{1} \\ g_{2} \\ g_{3} \end{bmatrix}$$
$$= \begin{bmatrix} \cosh(2b) & \sinh(2b) & 0 & 0 \\ \sinh(2b) & \cosh(2b) & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} g_{0} \\ g_{1} \\ g_{2} \\ g_{3} \end{bmatrix}$$
(1.176)

There are two important consequences of this result:

1. The matrix [M] is always real.

The elements of [M] are always real, despite the fact that the singular values of [S] are complex. This we can show in two stages: in the first case equation (1.175) generalises to transformation by the full Hermitian component of the polar decomposition [H], as shown in equation (1.177):

$$[R] = \frac{1}{2}[H] \cdot \begin{bmatrix} g_0 + g_1 & g_2 - ig_3 \\ g_2 + ig_3 & g_0 - g_1 \end{bmatrix} \cdot [H]^{*T} \Rightarrow [M] \quad m_{ij} \in \mathfrak{R}$$
(1.177)

The left-hand side is then always itself Hermitian, and so the Pauli expansion coefficients are real and hence all the elements of [M] must be real. Finally, by adding the unitary matrix component of the polar decomposition [U] we obtain a spin transformation of the form shown in equation (1.178):

$$[R] = \frac{1}{2} [U] \cdot [H] \cdot \begin{bmatrix} g_0 + g_1 & g_2 - ig_3 \\ g_2 + ig_3 & g_0 - g_1 \end{bmatrix} \cdot [H]^{*T} [U]^{*T}$$
$$\Rightarrow [M]' = \begin{bmatrix} 1 & 0 \\ 0^T & [O_3] \end{bmatrix} \cdot [M] \cdot \begin{bmatrix} 1 & 0 \\ 0^T & [O_3]^T \end{bmatrix}$$
(1.178)

However, we have already shown that the effect of a complex special unitary transformation maps into a real 3×3 rotation matrix $[O_3]$ that acts on the Cartesian components of the Stokes vector (equation (1.127)). Hence the unitary matrix component transforms [M], as shown in equation (1.178), where



Fig. 1.26 Minkowski diagram for canonical transformation of the Stokes vector

 $\underline{0} = (0, 0, 0)$ is a null vector. However, [M] remains real, and hence we have shown that for an arbitrary *complex* scattering matrix [S], there exists a corresponding 4×4 real matrix [M] which transforms the Stokes vectors as $\underline{g}_{new} = [M] \underline{g}$. This will be developed further in Section 2.2 concerning the Mueller matrix.

2. The matrix [M] represents a Lorentz 'boost' in the direction of the XPOL nulls.

Although the scattering amplitude matrix can always be made diagonal by using SVD, we have shown in equation (1.176) that the [M] matrix is never diagonal. We see that even in its most canonical form it has two off-diagonal elements. We now give a geometrical interpretation of this as a Lorentz transformation of the Stokes vector.

The transformation in equation (1.176) does not represent a plane rotation, but a Lorentz 'boost' in the subspace spanned by g_0 and g_1 . This can be made clear by reference to Figure 1.26. Here we show the g_0,g_1 plane. The lines at ± 45 degrees represent the loci of all possible states P and Q—the left quadrant being Q polarised and the right P polarised. These two lines form the analogue of the light cone in special relativity. They represent pure states of wave polarisation, but with varying amplitude. Note that points which lie below the lines represent the condition $g_1 > g_0$. As g_0 represents the total intensity of the wave it is not possible to satisfy this condition for plane waves. Hence this region is nonphysical, and corresponds to light speeds greater than *c* in special relativity. But what about the region above the lines? Here we have the condition $g_0 > g_1$. This can be allowed if we choose, for example, some subspace of the signal to represent g_1 . This occurs with wave depolarisation arising from partial coherence between the states (see Chapter 2).

In the extreme case $g_1 = 0$, and we have noise signals represented by the vertical axis and showing no preference for P or Q polarisation. We note that the geometric effect of [*M*] on the coordinate system is to distort it by an angle δ as shown. Hence the set of incident waves defined by a g_1/g_0 ratio lying along the line at δ degrees are transformed on scattering into waves with $g_1^{new} = 0$; that is, into noise and *vice versa*. Borrowing terminology from special relativity this represents a Lorentz 'boost' in the direction PQ in polarisation space; that is, the space of the Poincaré sphere and not physical space. The magnitude of the 'boost' depends only on the ratio of singular values of [*S*] or fork angle γ such that the angle δ is given by

$$\tan \delta = \tanh 2b = \frac{1 - \tan^4 \gamma}{1 + \tan^4 \gamma} \tag{1.179}$$

Geometrically, therefore, the pure states P and Q remain unchanged (they are the XPOL nulls of the scattering matrix and so by definition are invariant). This corresponds directly to the invariance of the speed of light in special relativity. However, the coordinates of other points do change, and we shall see in Chapter 2 that this has important physical interpretation in terms of depolarisation.

In summary we have seen that we can extend the idea of transformation of polarisation basis by a unitary matrix to transformation of polarisation state by scattering. In the unitary change of base we can write the general change of base as a matrix exponential function (equation (1.180)), exposing as it does the elementary geometry of the Poincaré sphere:

$$\underline{E}' = [U_2]\underline{E} = \exp(-i\theta\underline{n} \cdot \underline{\sigma})\underline{E} = (\cos\theta\sigma_0 - i\sin\theta\underline{n} \cdot \underline{\sigma})\underline{E}$$
(1.180)

where $\underline{\sigma}$ are the set of three Pauli matrices, and \underline{n} is a real three-vector defining the axis for rotation through θ in R3—a real three-dimensional space of the Stokes vector. The most general form of transformation of \underline{E} that accounts not only for change of base but also polarisation dependent scattering, can be written in the form shown in equation (1.181):

$$\underline{E}' = [S] \underline{E} = \exp((\delta p - i\theta \underline{n}) \cdot \underline{\sigma}) \underline{E}$$
(1.181)

where δ, θ are real, and <u>p</u> and <u>n</u> are three-vectors. For the important case of symmetric scattering matrices (as occurs for backscatter in the sensor coordinates, for example) it follows that <u>n</u> = (1,0,0), in which case the most general scattering matrix may be written in geometrical form, as shown in equation (1.182):

$$[S] = \exp((\delta \cos a - i\theta)\sigma_x + \mu \cos b\sigma_y + \mu \cos c\sigma_z)$$
(1.182)

which corresponds to a Lorentz boost of magnitude δ in a direction in Stokes space given by Euler angles *a*, *b* and *c*, defining the fork axis PQ in Figure 1.25.

1.6 The scattering vector formulation

In the previous section we were concerned with an interpretation of the form of the scattering matrix based on the extreme points of the amplitude scattering function. We saw, however, that the scattering matrix can be characterized using an alternative approach based on its transformation properties. In this section we start from this idea to develop the scattering vector characterization (Byrne, 1971; Cloude, 1985). This will lead us to consider several important new ideas such as orthogonality of matrices and generalized rotation invariance, and lead us to develop a new method for characterising scattering matrices in general. The vectorization of a matrix is the straightforward process of expanding the matrix using a set of simpler basis elements. The coefficients of this expansion then form a vector representation. To start with a simple but important example, we consider a straightforward lexicographic expansion as shown in equation (1.183):

$$[S] = \begin{bmatrix} x & y \\ w & z \end{bmatrix} = x \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} + y \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} + w \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} + z \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$$
(1.183)

One way to consider this expansion is as a representation of the information in the scattering matrix in terms of a basic set of canonical scattering mechanisms, each represented by the simpler matrices shown on the right-hand side. In this regard the basis elements for the lexicographic expansion have a particular physical significance as electric dipole and dihedral scatterers, 'x' as horizontal, 'z' as vertical and the combination of y+w as a 45-degree dihedral (see equation (1.161)). With this idea in mind, we have seen that the set of Pauli spin matrices have important canonical interpretation in terms of generic specular and dihedral scattering. This motivates the following so-called Pauli expansion of the [S] matrix as follows:

$$[S] = \frac{1}{\sqrt{2}} \begin{bmatrix} a+b & c-id\\ c+id & a-b \end{bmatrix}$$
$$= \frac{a}{\sqrt{2}} \begin{bmatrix} 1 & 0\\ 0 & 1 \end{bmatrix} + \frac{b}{\sqrt{2}} \begin{bmatrix} 1 & 0\\ 0 & -1 \end{bmatrix} + \frac{c}{\sqrt{2}} \begin{bmatrix} 0 & 1\\ 1 & 0 \end{bmatrix} + \frac{d}{\sqrt{2}} \begin{bmatrix} 0 & -i\\ i & 0 \end{bmatrix}$$
(1.184)

where the square-root factor is used to keep the total scattered power constant; that is, $|x|^2 + |y|^2 + |w|^2 + |z|^2 = |a|^2 + |b|^2 + |c|^2 + |d|^2$. Note that we can relate the vectors in the two systems—lexicographic and Pauli—by a 4 × 4 unitary matrix transformation as shown in equation (1.185):

$$\begin{bmatrix} x \\ y \\ w \\ z \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & -i \\ 0 & 0 & 1 & i \\ 1 & -1 & 0 & 0 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix}$$
(1.185)

It is then straightforward to see how we may generalize this idea of expansion to include any set of four 'complete' matrices ('complete' in the sense that any 2×2 complex matrix must be able to be represented by the set). In general we can formalize the vectorization of [S] by defining a Pauli scattering vector <u>k</u> as shown in equation (1.186):

$$[S] = \begin{bmatrix} S_{XX} & S_{XY} \\ S_{YX} & S_{YY} \end{bmatrix} \Rightarrow \underline{k} = \frac{1}{\sqrt{2}} \begin{bmatrix} S_{XX} + S_{YY} \\ S_{XX} - S_{YY} \\ S_{XY} + S_{YX} \\ i(S_{XY} - S_{YX}) \end{bmatrix} = \begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix}$$
$$= |\underline{k}|\underline{w}, |\underline{k}| = \sqrt{aa^* + bb^* + cc^* + dd^*}$$
(1.186)

and then representing the effect of using a basis set other than the Pauli matrices by a 4×4 unitary matrix transformation of the Pauli coefficients

(a generalization of equation (1.185)), as shown in equation (1.187):

$$\underline{k}' = [U_4]\underline{k} \tag{1.187}$$

Note, as a special case, that the effect of change of coordinates from wave (FSA) to sensor (BSA) on the Pauli vector can be represented by such a unitary matrix, as shown in equation (1.188):

$$[S]_{FSA} = \begin{bmatrix} x & y \\ w & z \end{bmatrix} \Rightarrow [S]_{BSA} = \begin{bmatrix} x & y \\ -w & -z \end{bmatrix}$$
$$\Rightarrow \underline{k}_{BSA} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -i \\ 0 & 0 & i & 0 \end{bmatrix} \underline{k}_{FSA}$$
(1.188)

1.6.1 Scattering mechanisms

As shown in equation (1.186), it is always possible to normalize the scattering vector by its amplitude to generate a complex unitary vector w. This normalized vector w we call a 'scattering mechanism', as it can be used to characterize differences in polarised wave scattering (see Chapter 3). In particular it is the transformation properties of this vector that we will now study. Note that the concept of a scattering mechanism conveniently scales with dimension of the problem. The only constraint is that w must always have unit amplitude. We can start with N = 1; that is, scalar fields characterized by a single complex number. In this case w is a phase term $e^{i\phi}$; that is, the set of transformations are those associated with phase shifts of the scattered field. In two dimensions we can combine phase shifts with amplitude ratios and still keep w unitary, as shown in equation (1.189). A physical interpretation of the two parameters α_w and $\phi_2 - \phi_1$ in terms of coordinates on the Poincaré sphere was developed for polarised waves in Section 1.3. Key for our discussion in terms of the scattering matrix is the extension of this idea to three, four and higher dimensions. These can be constructed by extension as shown in equation (1.189):

$$\underline{w}_{1} = e^{i\phi} \quad \underline{w}_{2} = \begin{pmatrix} \cos \alpha_{w} e^{i\phi_{1}} \\ \sin \alpha_{w} e^{i\phi_{2}} \end{pmatrix}$$
$$\underline{w}_{3} = \begin{pmatrix} \cos \alpha e^{i\phi_{1}} \\ \sin \alpha \cos \psi e^{i\phi_{2}} \\ \sin \alpha \sin \psi e^{i\phi_{3}} \end{pmatrix} \quad \underline{w}_{4} = \begin{pmatrix} \cos \alpha e^{i\phi_{1}} \\ \sin \alpha \cos \psi e^{i\phi_{2}} \\ \sin \alpha \sin \psi \cos \gamma e^{i\phi_{3}} \\ \sin \alpha \sin \psi \sin \gamma e^{i\phi_{4}} \end{pmatrix}$$
(1.189)

As we did for wave states using the Poincaré sphere, we now need to develop a physical interpretation of the parameters involved in these higher dimensional vectors. We begin by recognising that in general, any pair of sets of four basis matrices for the expansion of [S] can be related in a smooth and continuous way. To move from one set to another we employ unitary transformations (see Appendix 2 for a discussion of general unitary transformations). Mathematically we can then relate the vector of coefficients in one basis \underline{k} to those in the second basis set $\underline{k'}$ by a unitary matrix transformation, itself expressed as a

matrix exponential function, so forming a natural multidimensional extension of the simple scalar case, as shown in equation (1.190):

$$\underline{k}' = [U_N]\underline{k} \rightarrow \begin{cases} [U_1] = e^{i\phi} \\ [U_2] = e^{i\phi\underline{n}.\underline{P}} \\ [U_3] = e^{i\phi\underline{n}.\underline{G}} \\ [U_4] = e^{i\phi\underline{n}.\underline{D}} \end{cases}$$
(1.190)

Here we have used the notation P for the set of three Pauli matrices, G for the set of eight Gell–Mann matrices, and D the sixteen Dirac matrices (see Appendix 2 for definitions of these matrix sets). These compact representations will be useful when we turn to consider depolarisation effects in Chapter 2.

Depolarisation and scattering entropy

2

This chapter is concerned with the process of wave depolarisation and its formal description using the algebraic and geometrical tools developed in Chapter 1. Depolarisation is inherently a stochastic process, and somewhat destroys the transfer of vector information from source to far field. Methods for quantifying this loss of information will be developed based on the concept of scattering entropy, which we shall see provides a unifying concept of depolarisation across many types of polarisation problems. In Appendix 3 we outline the basic features of multivariate statistical signal analysis required for a description of coherence and its relation to depolarisation in polarimetric studies. Readers not familiar with these basic stochastic concepts should consult the Appendix for more details.

The first important point is to draw a distinction between wave depolarisation and crosspolarisation. These are often confused in the literature, and yet have important and subtle distinguishing characteristics, as we now demonstrate.

The term 'crosspolarisation' refers to any process in wave propagation and scattering that causes coupling between orthogonal states of polarisation. This process can be deterministic or stochastic in nature. For example, backscatter from a dipole oriented at θ degrees has a scattering matrix as shown in equation (2.1). Here we can see a level of crosspolarisation given by $\cos \theta \sin \theta$. As a second example, consider wave propagation though a half-wave plate (that is, a lossless birefringent plate with a thickness that generates a 180-degree phase shift between its eigenpolarisations at the design wavelength λ). When such a plate is oriented at θ degrees to the wave coordinates, the Jones matrix has the



Fig. 2.1 The set of depolarising systems as a subset of cross-polarisation

form shown in equation (2.2).

$$[S] = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix}$$
$$= \begin{bmatrix} \cos^{2}\theta & \cos\theta\sin\theta \\ \cos\theta\sin\theta & \sin^{2}\theta \end{bmatrix}$$
$$[M_{2}] = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix}$$
$$= \begin{bmatrix} \cos^{2}\theta - \sin^{2}\theta & 2\cos\theta\sin\theta \\ 2\cos\theta\sin\theta & \sin^{2}\theta - \cos^{2}\theta \end{bmatrix}$$
(2.2)

Here again we see a coupling of energy between modes, this time given by $\sin 2\theta$. Such an element is widely used, for example, to rotate the polarisation of a laser beam.

The key feature relating these two examples is that there exists some *deterministic* matrix transformation that can be used to *remove* the crosspolarisation. Hence neither of these systems depolarise the incident wave.

Depolarisation is therefore a coupling of energy from deterministic into stochastic modes of the field. It is connected with reversibility of crosspolarisation processes. It can be considered a noise-generating process, although as we shall show, information can still be extracted from the stochastic modes. Hence as shown schematically in Figure 2.1, all depolarising systems cause crosspolarisation but not *vice versa*. This chapter is concerned with systems that lie in the shaded region of Figure 2.1.

Note that here we define a depolariser as a system that causes noise generation when illuminated by a pure polarised wave (waves of the form shown in equation (1.133)). This avoids any of the semantic issues raised, for example, in Hovenier (2004) concerning the definition of depolarisation when considering illumination by partially polarised waves. We have already seen in Section 1.5.3 that the Lorentz transformation makes it possible even for simple point scatterers, characterized by a single amplitude matrix [S], to both increase and decrease the coherence of partially polarised waves according to the geometry of the Lorentz boost. However, the conservation of zero wave entropy means that such systems are here characterized as polarising. We now turn to consider systems where such a conservation law no longer applies.

2.1 The wave coherency matrix

In Figure 1.11 and accompanying discussion of polarisation geometry we established that even though plane waves are dynamic in nature, the geometrical parameters of the polarisation ellipse are invariant with time. However, in scattering and propagation through random media the geometry of the ellipse can become a function of space/time due, for example, to motion of particles in the scattering medium or to coherent fluctuations associated with speckle. Under such circumstances the ellipse itself becomes a dynamic quantity. It is then of interest to see if there exists some set of average parameters that can be used, together with some measure of spread or variance, to describe the wave state. Note that here we limit attention to quasimonochromatic waves—those for which the spectrum of fluctuations $\Delta \omega$ is small compared to ω . For an introduction to the topic of wideband polarimetry when this is no longer true, see Nelander (1995).

To quantify the stochastic nature of wave depolarisation, we first select an orthogonal polarisation base x,y, and then form averages over all possible products of complex field components. This is achieved by defining a 2 × 2 wave coherency matrix $\langle [J] \rangle$ (Born and Wolf, 1989) from the outer product of polarisation vectors, as shown in equation (2.3), where from equation (A3.9) it follows that $|j_{12}|^2 \leq |j_{11}| \cdot |j_{22}|$ and so det([J]) ≥ 0 .

$$[J] = \langle \underline{E} \cdot \underline{E}^{*T} \rangle = \begin{bmatrix} \langle E_x E_x^* \rangle & \langle E_x E_y^* \rangle \\ \langle E_y E_x^* \rangle & \langle E_y E_y^* \rangle \end{bmatrix} = \begin{bmatrix} j_{11} & j_{12} \\ j_{12}^* & j_{22} \end{bmatrix} = [J]^{*T}$$
(2.3)

This provides an important link with Lorentz geometry via the Minkowski metric, as discussed in Chapter 1. Hence [J] is a positive semidefinite (PSD) Hermitian matrix; that is, its eigenvalues are both real and non-negative. It has an associated real quadratic form defined as shown in equation (2.4):

$$\underline{w}^{*T}\left[J\right]\underline{w} \ge 0 \tag{2.4}$$

where \underline{w} is a two-element unitary complex wave vector. We can now use this matrix to obtain a general expression for the coherence between signals in arbitrary p and q polarisation channels. We start by defining the two desired receiver channels on the Poincaré sphere from their unitary \underline{w} vectors, as shown in equation (2.5):

$$\underline{w}_{P} = \begin{bmatrix} \cos \alpha_{P} \\ \sin \alpha_{P} e^{i\delta_{P}} \end{bmatrix} \underline{w}_{Q} = \begin{bmatrix} \cos \alpha_{Q} \\ \sin \alpha_{Q} e^{i\delta_{Q}} \end{bmatrix}$$
(2.5)

The received signals in these components are then obtained by projection onto the incident \underline{E} vector. The 'PQ' coherence can then be obtained as shown in equation (2.6):

$$\begin{cases}
S_{P} = \underline{w}_{P}^{*T} \underline{E} \\
S_{Q} = \underline{w}_{Q}^{*T} \underline{E}
\end{cases} \Rightarrow \gamma_{PQ} = \frac{\langle S_{P} S_{Q}^{*} \rangle}{\sqrt{\langle S_{P} S_{P}^{*} \rangle \langle S_{Q} S_{Q}^{*} \rangle}} \\
\frac{\underline{w}_{P}^{*T} \langle \underline{E} . \underline{E}^{*T} \rangle \underline{w}_{Q}}{\sqrt{\underline{w}_{P}^{*T} \langle \underline{E} . \underline{E}^{*T} \rangle \underline{w}_{Q}}} = \frac{\underline{w}_{P}^{*T} [J] \underline{w}_{Q}}{\sqrt{\underline{w}_{P}^{*T} [J] \underline{w}_{P} . \underline{w}_{Q}^{*T} [J] \underline{w}_{Q}}}
\end{cases}$$
(2.6)

Hence [J] contains all the information required to synthesize the coherence for arbitrary polarisation channels p and q. For this reason it is termed the wave coherency matrix, and forms the central focus for a quantitative treatment of wave depolarisation. Note that if p = q then the coherence is always unity as expected (we then compare a signal channel with itself). Often p and q will be orthogonal, when for example we wish to measure the full polarisation state of a wave. In this case we note that if we choose an orthogonal w pair such that the following condition applies:

$$[J] \underline{w}_{O} = \lambda \underline{w}_{O} \tag{2.7}$$

—that is, as an eigenvector of [J]—then from equation (2.6) the coherence will always be zero. Hence the eigenvalue problem of [J] is linked to the coherence

properties of the wave. To pursue this idea further, we note that for a change of polarisation base we have $\underline{E}_2 = [U_2] \underline{E}$, and the new coherency matrix can then be easily derived as shown in equation (2.8):

$$[J]_{2} = \langle \underline{E}_{2} . \underline{\underline{E}}_{2}^{*} \rangle = \langle [U_{2}] \, \underline{\underline{E}} . \underline{\underline{E}}^{*T} \, [U_{2}]^{*T} \rangle$$
$$= [U_{2}] . \langle \underline{\underline{E}} . \underline{\underline{E}}^{*T} \rangle [U_{2}]^{-1} = [U_{2}] . [J] . [U_{2}]^{-1}$$
(2.8)

It follows from the properties of Hermitian matrices that we can then always find a base such that [J] is diagonal, as shown in equation (2.9):

$$[J] = [U_2] \cdot [D] \cdot [U_2]^{*T} = \begin{bmatrix} \underline{w} & \underline{w}_{\perp} \end{bmatrix} \cdot \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} \cdot \begin{bmatrix} \underline{w} & \underline{w}_{\perp} \end{bmatrix}^{*T} \quad \lambda_1 \ge \lambda_2 \ge 0$$
(2.9)

Here \underline{w} is the eigenvector corresponding to the maximum eigenvalue λ_1 , and $[U_2]$ is a 2×2 unitary matrix formed with columns generated by the (orthogonal) eigenvectors. The eigenvector is then associated with a state P on the Poincaré sphere. Note that this state is not uniquely defined, as any component of [U] obtained by exponentiation of the Cartan sub-algebra of SU(2) (Cornwell, 1984; Georgi, 1999) will be removed by the eigenvector decomposition of equation (2.9) (see Appendix 2). Since SU(2) has a one-dimensional Cartan sub-algebra, these 'hidden' variables are obtained for phase transformations of the form shown in equation (2.10):

$$[U]_{\text{cartan}} = \begin{bmatrix} e^{i\phi} & 0\\ 0 & e^{-i\phi} \end{bmatrix}$$
(2.10)

which, as we have shown in Section 1.3.2, corresponds to the absolute phase of the polarised wave. Hence the eigenvectors of [J] have an arbitrary phase, and equation (2.8) remains invariant to such changes. This is just a formal way of saying that the coherency matrix contains information only about relative phase angles or phase differences between channels. Hence [J] has four independent parameters: two real eigenvalues and two angles on the Poincaré sphere locating the point P. In our discussion of polarised plane waves we saw that three parameters were sufficient to represent a state of elliptical polarisation. Hence [J] contains one new parameter that is associated with loss of coherence or depolarisation. To establish a general notation for this type of parameter matching we anticipate generalization of [J] and write a general N × N coherency matrix in the following form (Cloude, 2001a):

$$[C] = (E + L) + [E + L]$$
(2.11)

where (..) are the polarised components, [..] the depolarising parameters, and for each, E are those parameters associated with the eigenvectors and L the eigenvalues. From the general structure of $N \times N$ coherency matrices we then have the following constraints (see Appendix 2):

- L + (L) = N
- E + (E) = dim(SU(N)) -rank(SU(N))

With this notation in place we can now write the wave coherency matrix in the following compact form:

$$[J] = (2+1) + [0+1]$$
(2.12)

It follows from this that in order to consider the effects of wave depolarisation (N = 2) we need only consider diagonal matrices, as the depolarisation parameter lies entirely in the eigenvalue spectrum and not in the eigenvectors. We shall see this is not true for higher dimensional problems. This special case of N = 2 then leads to a further important result known as the wave dichotomy.

2.1.1 The wave dichotomy

In a sense, the eigenvector decomposition states that for any wave we can find an orthogonal base for which the numerator in equation (2.6), and hence the coherence, is zero. Two special cases will illustrate extreme forms of this result. If the signal is noise-like then its coherency matrix will have the diagonal form shown in equation (2.13):

$$[J]_{noise} = \begin{bmatrix} \lambda & 0\\ 0 & \lambda \end{bmatrix}$$
(2.13)

In this case the coherence in equation (2.6) will be zero for arbitrary choice of orthogonal p and q; that is, the signal is indistinguishable from noise, there is no variation of coherence with polarisation, and depolarisation is complete. The whole signal is characterized by only a single parameter: the noise variance, λ . This represents the most extreme case of polarisation 'memory loss'.

At the other extreme, if we consider a pure polarised wave then det([J]) = 0 (see Section 1.3.4), and the matrix must then have one zero eigenvalue; that is, [J] is of the form shown in equation (2.14), where λ is in this case the squared amplitude of the polarisation ellipse:

$$[J]_{polarised} = \begin{bmatrix} \lambda & 0\\ 0 & 0 \end{bmatrix}$$
(2.14)

In this case the coherence is unity for all choices of p and q, except when p corresponds to the polarisation state itself. In this singular case the signal component in the orthogonal channel is by definition zero and hence the coherence function is no longer defined, and the numerator and denominator will now both be zero. In this case the signal has zero depolarisation and is purely deterministic in nature.

In general, of course, real signals lie between these two extremes, with a coherence that depends on the choice of polarisation base. These examples illustrate why the coherence is itself a bad way to characterize the depolarisation properties of the signal. Not only does it vary with polarisation state but it can also be singular for the case of pure polarised waves. We require instead a basis invariant regular parameter to account for the level of depolarisation in the signal. This can be defined in several ways—for example, from the ratio of eigenvalues—by defining the degree of polarisation D_p as shown in

equation (2.15):

$$D_p = \frac{\lambda_1 - \lambda_2}{\lambda_1 + \lambda_2} \quad 0 \le D_p \le 1 \tag{2.15}$$

This function is by definition invariant to changes of polarisation base, and so represents a fundamental property of the signal. Like coherence magnitude it lies between 0 and 1; when $D_p = 0$ we have a noise signal, while $D_p = 1$ for pure polarised waves.

This parameterization also allows us to derive a decomposition of the coherency matrix into 'polarised' and 'depolarised' components by writing an arbitrary [J] as the sum of a polarised wave plus noise, as shown in the top row of equation (2.16):

Wave Decomposition 1
$$[J] = [U] \cdot \begin{bmatrix} \lambda_1 - \lambda_2 & 0 \\ 0 & 0 \end{bmatrix} \cdot [U]^{*T} + \lambda_2 \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

Wave Decomposition 2 $[J] = \lambda_1 \underline{w} \cdot \underline{w}^{*T} + \lambda_2 \underline{w}_{\perp} \cdot \underline{w}_{\perp}^{*T}$

$$(2.16)$$

However, there is a second eigenvalue factorization that is equally valid. Instead of considering a noise + signal type decomposition, we can alternatively decompose the problem into an independent mixture of 'coherent states', as shown in the lower row of equation (2.16). This existence of a pair of statistical models for characterising depolarisation is known as the wave dichotomy. While the noise + signal approach led to D_p as a measure of depolarisation, the mixture of states approach leads to a second measure, namely the wave entropy H_w . Entropy in general is defined as the distribution of probabilities across a set of physical states (Wiener, 1930; Wolf, 1954; O'Neill, 1991; Brosseau, 1998). To calculate the entropy in this case we need to generate the probabilities of the two states in the model. As they are independent this is very easy, as the probability of each state is just given by its normalized eigenvalue, so that in general we can write the two probabilities of the orthogonal states as shown in equation (2.17).

$$P_i = \frac{\lambda_i}{\sum \lambda} \quad 0 \le P_i \le 1 \tag{2.17}$$

 P_i can be considered the probability that the state represented by the ith eigenvector will occur. With this definition in place, the entropy can then be defined as shown in equation (2.18):

$$H_w = -\sum_{i=1}^{2} p_i \log_2 p_i \quad 0 \le H_w \le 1$$
(2.18)

When the entropy is zero we have zero uncertainty as to the state of polarisation; that is, a coherence of one and a purely polarised wave. At the other extreme, when the entropy is one then we have maximum uncertainty as to the state of polarisation and a noise signal with a coherence of zero.

The solution to the wave dichotomy remains a matter largely of choice in wave problems, but as we shall see for higher order coherency matrices the entropy approach generalizes in a straightforward way while the signal + noise methods are plagued by ambiguities. This can be attributed to the special case of N = 2, which has all its depolarisation effects contained in the eigenvalue spectra (equation (2.12)).

Returning for a moment to the signal + noise approach and the degree of polarisation, one key feature of this method is its relationship to the Stokes vector representation of polarised waves (see Section 1.3.4). Even in the presence of depolarisation we can still expand the coherency matrix in terms of four Stokes parameters to obtain a Stokes vector as shown in equation (2.19):

$$[J] = \begin{bmatrix} \langle E_x E_x^* \rangle & \langle E_x E_y^* \rangle \\ \langle E_y E_x^* \rangle & \langle E_y E_y^* \rangle \end{bmatrix} = \frac{1}{2} \begin{bmatrix} g_0 + g_1 & g_2 - ig_3 \\ g_2 + ig_3 & g_0 - g_1 \end{bmatrix}$$
$$\Rightarrow \underline{g} = \begin{bmatrix} \langle E_x E_x^* \rangle + \langle E_y E_y^* \rangle \\ \langle E_x E_x^* \rangle - \langle E_y E_y^* \rangle \\ \langle E_x E_y^* \rangle + \langle E_y E_x^* \rangle \\ i \left(\langle E_x E_y^* \rangle - \langle E_y E_x^* \rangle \right) \end{bmatrix}$$
(2.19)

where now

$$\det([J]) = \lambda_1 \lambda_2 = \frac{1}{4} (g_0^2 - (g_1^2 + g_2^2 + g_3^2)) \ge 0$$

$$Trace([J]) = \lambda_1 + \lambda_2 = g_0 > 0$$
(2.20)

It follows from equations (2.15) and (2.20) that

$$\sum_{i=1}^{3} g_i^2 = (\lambda_1 - \lambda_2)^2 \Rightarrow D_p = \frac{\sqrt{\sum_{i=1}^{3} g_i^2}}{g_0}$$
(2.21)

and so the degree of polarisation can be written as a ratio of two scalars derived directly from the Stokes vector. The first—the g_0 element—is just the total wave intensity, while the second is the Euclidean norm of a three-vector in the space of the Poincaré sphere. This suggests a parallel decomposition of the Stokes vector into polarised and depolarised components, as shown in the top row of equation (2.22). Here the first component represents a polarised wave and a point on the Poincaré sphere. The second represents a noise signal, which has zero everywhere except in the first element. This corresponds directly to wave decomposition 1 in equation (2.16). The wave dichotomy means we can also write an arbitrary Stokes vector as the sum of two orthogonal states. The components of the two-state decomposition can be easily derived as shown in the lower portion of equation (2.22).

Wave Decomposition 1
$$\underline{g} = \begin{bmatrix} \sqrt{\sum_{i=1}^{3} g_i^2} \\ g_1 \\ g_2 \\ g_3 \end{bmatrix} + \begin{bmatrix} g_0 - \sqrt{\sum_{i=1}^{3} g_i^2} \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

Wave Decomposition 2
$$\begin{cases} g_{01} = \frac{g_0 + \sqrt{\sum_{i=1}^{3} g_i^2}}{2} \\ g_{02} = \frac{g_0 - \sqrt{\sum_{i=1}^{3} g_i^2}}{2} \\ g_{02} = \frac{g_0 - \sqrt{\sum_{i=1}^{3} g_i^2}}{2} \end{cases}$$
$$\Rightarrow \underline{g} = g_{01} \begin{bmatrix} \frac{1}{\frac{g_1}{\sqrt{\sum_{i=1}^{3} g_i^2}}} \\ \frac{g_2}{\sqrt{\sum_{i=1}^{3} g_i^2}} \\ \frac{g_2}{\sqrt{\sum_{i=1}^{3} g_i^2}} \end{bmatrix} + g_{02} \begin{bmatrix} \frac{1}{\frac{-g_1}{\sqrt{\sum_{i=1}^{3} g_i^2}}} \\ \frac{-g_2}{\sqrt{\sum_{i=1}^{3} g_i^2}} \\ \frac{-g_3}{\sqrt{\sum_{i=1}^{3} g_i^2}} \end{bmatrix}$$
(2.22)

In summary, we have seen that there are several ways to characterize depolarised waves. Firstly we note that while depolarisation is associated with a loss of coherence, it is better to characterize it on the basis of the eigenvalue spectrum of the wave coherency matrix rather than in terms of the coherence function itself. Arising from this is the wave dichotomy, whereby the eigenvalues can be used either as part of a signal + noise model which leads to the degree of polarisation, or as a mixture of states model which leads to the definition of wave entropy. Both have a simple representation in terms of the Stokes vector.

Entropy can be viewed as an information destroying process, and with this in mind it is interesting to characterize the various ways in which entropy can be increased or decreased by scattering. The simplest way to increase entropy is to consider the addition of thermal noise to a signal. Entropy can also be increased at source, as is the case for passive remote sensing when the source is thermal radiation. However, for active sensors the source usually has zero entropy by design. Consequently, any change in entropy must be caused by environmental scattering and propagation. We now turn to consider a detailed analysis of such effects.

2.2 The Mueller matrix

The starting point for an analysis of the effects of wave scattering on depolarisation is to consider the wave coherency matrix obtained *after* wave scattering $[J_s]$ compared to that before, $[J_i]$. Using equation (2.3) and the definition of the scattering amplitude matrix [S] we obtain the following relationship between the two coherency matrices, where [S] is an arbitrary 2 × 2 complex amplitude matrix:

$$[J_s] = [S] . [J_i] . [S]^{*T}$$
(2.23)

Clearly [S] will have the effect of changing the eigenvalues and eigenvectors of [J], and hence will change the degree of wave depolarisation. By expanding the matrix product in equation (2.23) and collecting terms of [J] to form elements of the corresponding Stokes vectors \underline{g}_i and \underline{g}_s we can rewrite equation (2.23) as a linear transformation between vectors, as shown in equation (2.24):

$$\underline{g}_{s} = [M] \underline{g}_{i} \tag{2.24}$$

The matrix [M] is called the Mueller matrix, after Hans Mueller, an MIT physicist who first derived the properties of such matrices (but seems to have published very little in the open literature). Parallel development of a similar approach can be found in the contemporary work of Perrin (1942).

We have already developed a geometrical interpretation of equations of the form shown in equations (2.23) and (2.24). In equation (1.177) we showed that the elements of [M] are all real, despite the fact that [S] is complex. This can be confirmed by explicit derivation of the elements of [M] from [S], as we now demonstrate. We can reformulate equation (2.23) in terms of Stokes vectors by using the Kronecker matrix product (see Appendix 1), since for any set of three matrices we can write an expansion as shown in equation (2.25):

$$[A][X][B] \equiv [A] \otimes [B]^T \underline{x} \tag{2.25}$$

where *x* is a lexicographic row vectorization of the elements of [X]:

$$[X] = \begin{bmatrix} x_{11} & x_{12} & \dots & x_{1n} \\ x_{21} & x_{22} & \dots & x_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n1} & x_{n2} & \dots & x_{nn} \end{bmatrix} \Rightarrow \underline{x} = \begin{bmatrix} x_{11} \\ x_{12} \\ \vdots \\ x_{nn} \end{bmatrix}$$
(2.26)

Using this result, we can now write the vectorized wave coherency matrix of the scattered wave as shown in equation (2.27):

$$[S] = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \Rightarrow \underline{x}_{s} = [Y] \underline{x}_{i} = \begin{bmatrix} aa^{*} & ab^{*} & ba^{*} & bb^{*} \\ ac^{*} & ad^{*} & bc^{*} & bd^{*} \\ ca^{*} & cb^{*} & da^{*} & db^{*} \\ cc^{*} & cd^{*} & dc^{*} & dd^{*} \end{bmatrix} \cdot \begin{bmatrix} E_{x}E_{x}^{*} \\ E_{x}E_{y}^{*} \\ E_{y}E_{x}^{*} \\ E_{y}E_{y}^{*} \end{bmatrix}$$
(2.27)

We note that the properties of the scattered wave are determined by a 4×4 matrix of quadratic products of the elements of [S], arranged into a complex matrix [Y] as shown. In order to obtain the elements of the Mueller matrix [M], relating incident to scattered wave Stokes vectors, we now need to transform the vector <u>x</u> in equation (2.27) into the Pauli basis using the transformation

matrix shown in equation (2.28):

$$\underline{g} = \begin{bmatrix} E_x E_x^* + E_y E_y^* \\ E_x E_x^* - E_y E_y^* \\ E_x E_y^* + E_y E_x^* \\ i(E_x E_y^* - E_y E_x^*) \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & -1 \\ 0 & 1 & 1 & 0 \\ 0 & i & -i & 0 \end{bmatrix} \cdot \begin{bmatrix} E_x E_x^* \\ E_y E_x^* \\ E_y E_y^* \end{bmatrix}$$

$$\Rightarrow \begin{bmatrix} E_x E_x^* \\ E_x E_y^* \\ E_y E_x^* \\ E_y E_y^* \end{bmatrix} = \frac{1}{2} \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & i \\ 0 & 0 & 1 & -i \\ 1 & -1 & 0 & 0 \end{bmatrix} \cdot \begin{bmatrix} E_x E_x^* + E_y E_y^* \\ E_x E_x^* - E_y E_y^* \\ E_x E_y^* + E_y E_x^* \\ i(E_x E_y^* - E_y E_x^*) \end{bmatrix}$$
(2.28)

Combining equations (2.27) and (2.28) we finally obtain an expression for the matrix [M] in terms of the elements of [S] and [Y], as shown in equation (2.29):

$$[M] = \begin{bmatrix} m_{11} & m_{12} & m_{13} & m_{14} \\ m_{21} & m_{22} & m_{23} & m_{24} \\ m_{31} & m_{32} & m_{33} & m_{34} \\ m_{41} & m_{42} & m_{43} & m_{44} \end{bmatrix} \qquad m_{ij} \in \Re$$
$$= \frac{1}{2} \cdot \begin{bmatrix} 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & -1 \\ 0 & 1 & 1 & 0 \\ 0 & i & -i & 0 \end{bmatrix} \cdot \begin{bmatrix} aa^* & ab^* & ba^* & bb^* \\ ac^* & ad^* & bc^* & bd^* \\ ca^* & cb^* & da^* & db^* \\ cc^* & cd^* & dc^* & dd^* \end{bmatrix} \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & i \\ 0 & 0 & 1 & -i \\ 1 & -1 & 0 & 0 \end{bmatrix}$$
(2.29)

which can be expanded to yield an explicit mapping between any given [S] matrix and the sixteen real elements of [M], as shown in equation (2.30):

$$[M] = \frac{1}{2} \begin{bmatrix} aa^{*} + bb^{*} + cc^{*} + dd^{*} & aa^{*} - bb^{*} + cc^{*} - dd^{*} & 2Re(ab^{*} + cd^{*}) & 2Im(ab^{*} + cd^{*}) \\ aa^{*} + bb^{*} - cc^{*} - dd^{*} & aa^{*} - bb^{*} - cc^{*} + dd^{*} & 2Re(ab^{*} - cd^{*}) & 2Im(ab^{*} - cd^{*}) \\ 2Re(ac^{*} + bd^{*}) & 2Re(ac^{*} - bd^{*}) & 2Re(ad^{*} + bc^{*}) & 2Im(ad^{*} - bc^{*}) \\ -2Im(ac^{*} + bd^{*}) & -2Im(ac^{*} - bd^{*}) & -2Im(ad^{*} + bc^{*}) & 2Re(ad^{*} - bc^{*}) \end{bmatrix}$$

$$(2.30)$$

2.2.1 Properties of the Mueller matrix

Matrices of the type shown in equation (2.30)—those that can be written in the general form [M] = f([S])—have several interesting properties that distinguish them from general 4 × 4 real matrices (Abhyankar, 1969; Barakat, 1981, 1987; Kim, 1987; Cloude, 1989; Anderson, 1994; Hovenier, 1994, 2004). These are often referred to as 'pure' Mueller matrices in the literature. Here we summarize their key properties and highlight their significance in terms of the scattering of polarised waves (Mishchenko, 2000).

Mueller Property M1: if x = |det([S])| then $x^2 = m_{11}^2 - m_{21}^2 - m_{31}^2 - m_{41}^2$ **Mueller Property M2**: $|Tr([S])|^2 = Tr([M])$ so the trace of [M] is always non-negative

Mueller Property M3: x⁴ = det([M]) so the determinant of [M] can never be negative

Mueller Property M4: if $[S] \rightarrow [M]$ then $[S]^{-1} \rightarrow [M]^{-1}$ if $x \neq 0$

Mueller Property M5: $[S_1]$. $[S_2] \rightarrow [M_1]$. $[M_2]$

Mueller Property M6: $\alpha[S] \rightarrow |\alpha|^2[M] \ \alpha \in C$. This has important implications for the reversibility of the mapping from [S] to $[M], [S] = f^{-1}([M]),$ itself an important element in determining the level of depolarisation caused by scattering.

Mueller Property M7: $[S]^T \to [\Delta_4] [M]^T [\Delta_4] \qquad [\Delta_4] = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix}$ **Mueller Property M8**: $[S]^{*T} \rightarrow [M]^{T}$ Mueller Property M9: $[S]^* \to \Delta_4 [M] \Delta_4 \quad \Delta_4 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix}$

Mueller Property M10:

$$[S] = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \to [M] \Rightarrow \begin{bmatrix} a & -b \\ -c & d \end{bmatrix} \to [\Delta_{34}] [M] [\Delta_{34}]$$
$$[\Delta_{34}] = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix}$$

Mueller Property M11:

$$[S] = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \to [M] \Rightarrow \begin{bmatrix} d & b \\ c & a \end{bmatrix} \to [\Delta_2] [M]^T [\Delta_2]$$
$$[\Delta_2] = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Mueller Property M12:

$$[S]^{-1} = \frac{1}{\det([S])} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix} \rightarrow \frac{1}{x^2} [\Delta_{234}] [M]^T [\Delta_{234}]$$
$$[\Delta_{234}] = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix}$$

and so the inverse of [M] can be written as follows

$$[M]^{-1} = \frac{1}{x^2} [\Delta_{234}] [M]^T [\Delta_{234}]$$

Mueller Property M13:

$$[S] = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \to [M] \Rightarrow \begin{bmatrix} a & -c \\ -b & d \end{bmatrix} \to [\Delta_3] [M]^T [\Delta_3]$$
$$[\Delta_3] = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

The relation M13 is important for application of the reciprocity theorem to scattering problems (van de Hulst, 1981; Saxon, 1955). Property M12 in particular leads to two important further results, summarized in equation (2.31):

$$[M]^{-1} = \frac{1}{x^2} [\Delta_{234}] [M]^T [\Delta_{234}] \Rightarrow \begin{cases} [M]^T [\Delta_{234}] [M] = x^2 [\Delta_{234}] \\ Tr([M]^T [\Delta_{234}] [M]) = -2x^2 \end{cases}$$
(2.31)

These relations were first derived by Barakat (1981) and Simon (1987). Finally, an important simple and widely used relationship between the elements of the matrix [M], called the Fry–Kattawar relation (Fry, 1981) is shown in equation (2.32):

$$\sum_{i=1}^{4} \sum_{j=1}^{4} m_{ij}^2 = 4m_{11}^2$$
(2.32)

Although this relation is widely used in the literature, care must be taken in its application. For example, there exist matrices that satisfy this relation but do not have a corresponding single [S] matrix, and so for which the reverse mapping does not exist. A simple example is diag(1,1,1,-1). Hence equation (2.32) is a necessary but not sufficient condition for [M] to correspond to a single [S] (Hovenier, 1996, 2004).

All of these relations were developed with a view to establishing conditions for the reversibility of the relation $[S] = f^{-1}([M])$. In general, [M] = f([S])changes the degree of polarisation of the wave, but if the incident wave entropy is zero (a purely polarised wave) then the scattered wave entropy is also zero. This is termed the 'conservation of zero wave entropy', and is analogous to the conservation of light speed in special relativity (see Section 1.5.3). This property has to do with the question of reversibility of the mapping from [S] to [M]; that is, $[M] = f([S]) \Rightarrow [S] = f^{-1}([M])$? We can use property M6 to find the conditions for the inversion to be unique, since $z[S] \rightarrow |z|^2 [M]$ for all complex z, so $[S] \Rightarrow [M] \Rightarrow [S_R]$ is possible, but only where $[S_R]$ is the relative phase scattering matrix (that is, the scattering amplitude matrix can be reconstructed from [M], but only up to an arbitrary phase term). This helps establish conditions for uniqueness, but leaves the thorny issue of existence to be resolved. There is no guarantee that given a general 4×4 real matrix it will have a corresponding single [S] matrix representation. Before dealing with existence conditions, however, we turn to consider the special case of backscatter and the implications of the BSA coordinate system for the Mueller matrix.

2.2.2 The backscatter Mueller and Stokes reflection matrix

We saw in Chapter 1 that the case of backscatter is of particular interest when the structure of [S] and therefore [M] simplifies, as we now show. Following from Chapter 1, there are two versions of the backscatter [S] matrix, depending on the coordinate system employed (wave and sensor or FSA and BSA coordinates). Consequently there are two versions of [M] to consider. Equation (2.33) summarizes the backscatter form of [M] for the wave coordinate system (FSA). Equation (2.34) shows the corresponding form for the BSA coordinates. Note that in the BSA system the matrix is real symmetric. The BSA form of this matrix was first derived by Edward Kennaugh, at Ohio State University, in parallel with the optical developments by Mueller and Perrin, and so it is often termed the Kennaugh matrix, or sometimes simply the Stokes reflection matrix, in the radar literature (Kennaugh, 1952).

In both cases we see that the sixteen elements of [M] are reduced via the reciprocity theorem for backscatter to only ten (actually nine; see equation (2.38)). As [S] has only a maximum of six independent elements for backscatter, then clearly the elements of [M] are not all independent. Hence for a 4×4 real matrix [M] to correspond to a complex amplitude matrix [S]—to guarantee existence of the inverse mapping $[S] = f^{-1}([M])$ —the elements of [M] must satisfy some additional constraint equations:

$$\begin{split} [S] &= \begin{bmatrix} a & b \\ -b & d \end{bmatrix} \\ \Rightarrow [M] &= \frac{1}{2} \begin{bmatrix} aa^* + 2bb^* + dd^* & aa^* - dd^* & 2Re(ab^* - bd^*) & 2Im(ab^* - bd^*) \\ aa^* - dd^* & aa^* - 2bb^* + dd^* & 2Re(ab^* + bd^*) & 2Im(ab^* + bd^*) \\ 2Re(bd^* - ab^*) & 2Re(-ab^* - bd^*) & 2Re(ad^* - bb^*) & 2Im(ad^*) \\ 2Im(ab^* - bd^*) & 2Im(ab^* + bd^*) & -2Im(ad^*) & 2Re(ad^* + bb^*) \end{bmatrix} \\ &= \begin{bmatrix} m_{11} & m_{12} & m_{13} & m_{14} \\ m_{12} & m_{22} & m_{23} & m_{24} \\ -m_{13} & -m_{23} & m_{33} & m_{34} \\ m_{14} & m_{24} & -m_{34} & m_{44} \end{bmatrix} \end{split}$$
(2.33)
$$\begin{split} [S] &= \begin{bmatrix} a & b \\ b & -d \end{bmatrix}^* \\ &\Rightarrow [M] &= \frac{1}{2} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix} \\ &\cdot \begin{bmatrix} aa^* + 2bb^* + dd^* & aa^* - dd^* & 2Re(ab^* - bd^*) & 2Im(ab^* - bd^*) \\ aa^* - dd^* & aa^* - 2bb^* + dd^* & 2Re(ab^* + bd^*) & 2Im(ab^* + bd^*) \\ 2Re(ab^* - bd^*) & 2Re(ab^* + bd^*) & 2Re(bb^* - ad^*) & 2Im(ab^* + bd^*) \\ -2Im(ab^* - bd^*) & -2Im(ab^* + bd^*) & -2Im(-ad^*) & -2Re(ad^* + bc^*) \end{bmatrix} \\ &= \begin{bmatrix} m_{11} & m_{12} & m_{13} & m_{14} \\ m_{12} & m_{22} & m_{23} & m_{24} \\ m_{13} & m_{23} & m_{33} & m_{34} \\ m_{14} & m_{24} & m_{34} & m_{44} \end{bmatrix} \end{split}$$

Furthermore, there exists the possibility of formulating a set of Mueller matrices that do not correspond to a single [S] matrix at all. The most extreme example of these is the isotropic depolariser, with a matrix of the form shown in equation (2.35):

This matrix converts all Stokes vectors into a randomly polarised wave, hence its name. However, there is no corresponding single [S] matrix. Clearly, therefore, our analysis of depolarisation must go beyond relations like equation (2.32) and look more closely at the structure of Mueller matrices in general.

The set of Mueller matrices [M] form a subset of all possible 4×4 real matrices. Hence not all real matrices correspond to Mueller matrices and, because of depolarisation like that shown in equation (2.35), of those that do there is no guarantee that the inverse mapping to a single [S] will exist. This situation is summarized graphically in Figure 2.2. It is of practical interest, therefore, to determine two sets of conditions on a given matrix [M]—firstly to test if it corresponds to a Mueller matrix at all; and secondly, if it does, whether there is a corresponding single amplitude matrix [S]. By doing this we may then establish existence of the inverse mapping, and at the same time classify the different possible types of depolarisation. A full analysis of this problem follows Section 2.5, but here we make some general observations about the related topic of the Stokes criterion.

2.2.3 The Stokes criterion

An important set of constraints follows directly from equation (2.24) and the requirement that both sides must correspond to physical Stokes vectors, even in the presence of depolarisation. The so-called 'Stokes criterion' has been studied by many authors (Barakat, 1981, 1987; Fry, 1981; Gil, 1985, 1986; Cloude, 1989; Girgel, 1991; van der Mee, 1992, 1993; Givens, 1993; Hovenier, 1994, 1996), and defines valid Mueller matrices such that they satisfy the set of general mathematical conditions shown in equation (2.36):

$$\underline{g}_{s} = [M] \underline{g}_{i} \Rightarrow \begin{cases} g_{0s} \ge 0, g_{0i} \ge 0\\ g_{0s}^{2} - g_{1s}^{2} - g_{2s}^{2} - g_{3s}^{2} \ge 0\\ g_{0i}^{2} - g_{1i}^{2} - g_{2i}^{2} - g_{3i}^{2} \ge 0 \end{cases}$$
(2.36)

We can, for example, enforce these by insisting that the following quadratic form be positive semidefinite:

$$\underline{g}_{i}^{T} \left[\Delta_{234} \right] \left[M \right]^{T} \left[\Delta_{234} \right] \left[M \right] \underline{g} \ge 0 \tag{2.37}$$

where $[\Delta_{234}]$ is defined in property M12. This set of conditions, first derived in van der Mee (1992, 1993) and Givens (1993) in turn requires the eigenvalues of the composite matrix $[\Delta_{234}][M]^T [\Delta_{234}][M]$ to be real and non-negative. This can then be directly taken as a test of validity for candidate Mueller matrices.



A – the set of 4×4 real matrices

B - the set of Mueller matrices

C – the set of Mueller matrices for which $[S] = f^{-1}([M])$

Fig. 2.2 The set of Mueller matrices as a subset of real 4×4 matrices

However, this is a rather complicated and, as we shall see, incomplete formulation, and in the next section we derive a much simpler set of tests based on an eigenvalue analysis of the scattering coherency matrix.

Before leaving the Mueller matrix completely, however, we return again to the case of backscatter in equations (2.33) and (2.34). These equations give the impression that the backscatter matrix has ten independent elements. However, this is not true, there being only nine independent elements for the general case. The missing constraint equation arises again from the reciprocity theorem, which constrains the diagonal elements. In order for a general 4×4 real matrix to represent a backscatter Mueller matrix it must not only have the form shown in equations (2.33) and (2.34), but its diagonal elements must also satisfy the following trace conditions (Mishchenko, 1995):

Reciprocity
$$\Rightarrow \begin{cases} m_{11} - m_{22} - m_{33} - m_{44} = 0 & (BSA \text{ coordinates}) \\ m_{11} - m_{22} + m_{33} - m_{44} = 0 & (FSA \text{ coordinates}) \end{cases}$$
 (2.38)

The reason for this, as shown in the next section, is that the reciprocity theorem constrains an eigenvalue of the scattering coherency matrix to be zero. In this sense the coherency matrix formulation makes for a much simpler analysis of depolarisation properties, and so we now turn to consider its properties in more detail.

2.3 The scattering coherency matrix formulation

In the previous section we saw that the study of depolarisation is fundamentally related to a matrix of second order products of the scattering matrix, formed into a matrix [Y] as shown in equation (2.39):

$$[S] = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \Rightarrow \underline{x}_s = [Y] \ \underline{x}_i = \begin{bmatrix} a [S]^* & b [S]^* \\ c [S]^* & d [S]^* \end{bmatrix} \cdot \begin{bmatrix} E_x E_x^* \\ E_x E_y^* \\ E_y E_x^* \\ E_y E_y^* \end{bmatrix}$$
(2.39)

There are, however, two alternative ways of forming a matrix of all possible second-order products: the covariance and coherency matrices defined as outer products of the corresponding scattering vectors in the lexicographic and Pauli bases (see Section 1.6), as shown in equation (2.40) (Cloude, 1985, 1986). These matrices have a useful symmetry as, unlike [*Y*], they are complex Hermitian, having real elements along the diagonal and complex conjugate entries in symmetric off-diagonal positions. This reordering of the product terms has a more direct relationship to coherence between the elements of the <u>k</u> vector, hence the name coherency matrix. For example, [*C*] can be used to generalize the arguments about optimum polarisations treated for [*S*] matrices in Section 1.4.3 to depolarising scenarios (see Tragl (1990)). Furthermore, we can exploit the Hermitian symmetry by noting that all the 2×2 principal minors of [*C*] and [*T*] are zero, so for example, in equation (2.40) we have aa*bb* – ab*ba* = aa*cc* – ac*ca* = 0 and so on.

For a 4×4 matrix there are nine such minors, and for a 3×3 matrix there are four. In fact these provide just the set of required constraint equations between

the Mueller matrix elements discussed around equation (2.36). Rather than consider these in any more detail (to investigate further see Huynen (1970, 1987) and Hovenier (1994, 1996, 2004)) we proceed to the next key step and consider the effect of depolarisation on the form of the coherency matrix.

$$[S] = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \Rightarrow [C] = \underline{k}_{L} \underline{k}_{L}^{*T} = \begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix} \cdot \begin{bmatrix} a^{*} & b^{*} & c^{*} & d^{*} \end{bmatrix}$$
$$= \begin{bmatrix} a^{*} & ab^{*} & ac^{*} & ad^{*} \\ ba^{*} & bb^{*} & bc^{*} & bd^{*} \\ ca^{*} & cb^{*} & cc^{*} & cd^{*} \\ da^{*} & db^{*} & dc^{*} & dd^{*} \end{bmatrix} = [C]^{*T}$$
$$[S] = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \Rightarrow [T] = \underline{k}_{P} \underline{k}_{P}^{*T}$$
$$= \frac{1}{2} \begin{bmatrix} a+d \\ a-d \\ b+c \\ i(b-c) \end{bmatrix} \cdot [(a+d)^{*} \quad (a-d)^{*} \quad (b+c)^{*} \quad -i(b-c)^{*}] = [T]^{*T}$$
(2.40)

We can model depolarisation as fluctuations in space or time of the elements of the scattering amplitude matrix [S]. In general we then have the idea of a mean [S] matrix with fluctuations characterized as shown in equation (2.41), in which case we derive a corresponding statistical distribution of \underline{k} vectors:

$$[S(\underline{r},t)] = [\overline{S}] + [\Delta S(\underline{r},t)] \Rightarrow \underline{k}(\underline{r},t) = \underline{\overline{k}} + \Delta \underline{k}$$
(2.41)

The covariance and coherency matrices (and indeed the [Y] matrix) are then obtained as averages over the distribution, so, for example, over a population of L independent samples we can obtain an estimate of the mean coherency matrix as shown in equation (2.42):

$$\langle [T] \rangle = \sum_{i=1}^{L} \underline{k}_i \underline{k}_i^{*T}$$
(2.42)

In this case the principal minor relations, by the Schwarz inequality (see Appendix 3), now have the general form of non-negative inequalities, auch as $\langle aa^* \rangle \langle bb^* \rangle - \langle ab^* \rangle \langle ba^* \rangle \ge 0$. The Mueller matrix [*M*] corresponding to $\langle [T] \rangle$ is often called a sum-of-pure Mueller matrices (SPM) in the literature (Hovenier, 2004).

2.3.1 Eigenvalue decomposition of the coherency and covariance matrices

As [C] and [T] are formed from the sum of component Hermitian matrices (unlike [Y]), they remain positive semidefinite (PSD) Hermitian, which means

they maintain real non-negative eigenvalues and orthogonal eigenvectors, as shown in equation (2.43):

$$\langle [T] \rangle = [U_4] \cdot \begin{bmatrix} \lambda_1 & 0 & 0 & 0 \\ 0 & \lambda_2 & 0 & 0 \\ 0 & 0 & \lambda_3 & 0 \\ 0 & 0 & 0 & \lambda_4 \end{bmatrix} \cdot [U_4]^{*T} \implies \begin{cases} \lambda_1 \ge \lambda_2 \ge \lambda_3 \ge \lambda_4 \ge 0 \\ [U_4] \cdot [U_4]^{*T} = [I_4] \\ [U_4] = [\underline{e}_1 & \underline{e}_2 & \underline{e}_3 & \underline{e}_4] \\ (2.43) \end{cases}$$

where λ_i are the real eigenvalues, and \underline{e}_i the corresponding complex eigenvectors. Note that [*C*] has the same eigenvalues as [*T*] but its eigenvectors, \underline{e}_L , are related to those of [*T*], \underline{e}_P by a 4 × 4 unitary matrix, as shown in equation (2.44):

$$\underline{e}_{L} = \begin{bmatrix} U_{LP4} \end{bmatrix} \underline{e}_{P} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 & 0 & 0\\ 0 & 0 & 1 & -i\\ 0 & 0 & 1 & i\\ 1 & -1 & 0 & 0 \end{bmatrix} \underline{e}_{P} \Rightarrow \langle [C] \rangle = \begin{bmatrix} U_{LP4} \end{bmatrix} \langle [T] \rangle \begin{bmatrix} U_{LP4} \end{bmatrix}^{*T}$$
(2.44)

We have seen that for BSA reciprocal backscatter problems, the [S] matrix is symmetric, in which case the transformation from [T] to [C] is often reformulated as a 3×3 matrix, as shown in equation (2.45):

$$\stackrel{b=c}{\longrightarrow} \begin{bmatrix} a\\\sqrt{2}b\\d \end{bmatrix} = \begin{bmatrix} U_{LP3} \end{bmatrix} \frac{1}{\sqrt{2}} \begin{bmatrix} a+d\\a-d\\2b \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1&1&0\\0&0&\sqrt{2}\\1&-1&0 \end{bmatrix} \underline{e}_{P}$$

$$\Rightarrow \langle [C_3] \rangle = \begin{bmatrix} U_{LP3} \end{bmatrix} \langle [T_3] \rangle \begin{bmatrix} U_{LP3} \end{bmatrix}^{*T}$$

$$(2.45)$$

Note the factor of $\sqrt{2}$, which is required to maintain invariance of total scattered power. In what follows we refer to the coherency matrix [T], but understand that the equations could be easily recast in the form of the covariance matrix [C] using equation (2.44) or (2.45).

A key idea stemming from this eigenvalue decomposition is a generalization of the wave dichotomy to higher dimensional coherency matrices. In this case we can write the general depolarising coherency matrix as the sum of four independent and orthogonal scattering mechanisms, as shown in equation (2.46):

$$\langle [T] \rangle = \lambda_1 \underline{e}_1 \underline{e}_1^{*T} + \lambda_2 \underline{e}_2 \underline{e}_2^{*T} + \lambda_3 \underline{e}_3 \underline{e}_3^{*T} + \lambda_4 \underline{e}_4 \underline{e}_4^{*T}$$
(2.46)

where each vector corresponds to a relative phase amplitude matrix $[S_R]$ as shown in equation (2.47), and λ_i has a direct physical interpretation in terms of the power scattered into the mechanism represented by \underline{e} . For notational convenience we drop the brackets $\langle ... \rangle$, and simply refer to a general coherency matrix as [T] with a corresponding Mueller matrix [M]:

$$\underline{e} = \begin{bmatrix} e_a \\ e_b \\ e_c \\ e_d \end{bmatrix} \underline{e}^{*T} \underline{e} = 1 \Rightarrow [S_R] = \frac{1}{\sqrt{2}} \begin{bmatrix} e_a + e_b & e_c - ie_d \\ e_c + ie_d & e_a - e_b \end{bmatrix}$$
(2.47)



Fig. 2.3 The scattering triangle, relating the three depolarising matrices around the outside to [S] at the centre



While the formulation in terms of a 4×4 coherency matrix has several advantages, we note from equation (2.40) that [T] and [C] contain the same information on averages of second-order products as [Y]. Since [Y] is linked directly to the Mueller matrix [M] it follows that there exist 1–1 mappings between all three of these scattering matrices. The situation is summarized graphically in Figure 2.3. All descriptors of depolarisation stem from the amplitude matrix [S], and we have seen how to proceed from this starting point to each of the three matrices [M], [C] and [T] by averaging. What is missing, however, is how to go around the outside of this diagram; that is, to go from [M] to [T] directly without the need to go via [S]. The relationship between [T] and [C] has already been derived in equation (2.44). More significant is the relationship between [M] and [T], to which we now turn.

This can be derived by direct comparison between elements, and is given in detail in equation (2.48), with the inverse mapping provided for reference in equation (2.49):

$$\begin{split} [T] &= \begin{bmatrix} t_{11} & t_{12} & t_{13} & t_{14} \\ t_{12}^* & t_{23} & t_{33} & t_{34} \\ t_{14}^* & t_{24}^* & t_{34}^* & t_{44} \end{bmatrix} \Rightarrow [M] = \\ \\ \frac{1}{2} \begin{bmatrix} t_{11} + t_{22} + t_{33} + t_{44} & t_{12} + t_{12}^* - i(t_{34} - t_{34}^*) & t_{13} + t_{13}^* + i(t_{24} - t_{24}^*) & t_{14} + t_{14}^* - i(t_{23} - t_{23}^*) \\ t_{12} + t_{12}^* + i(t_{34} - t_{34}^*) & t_{11} + t_{22} - t_{33} - t_{44} & t_{23} + t_{23}^* + i(t_{14} - t_{14}^*) & t_{24} + t_{24}^* - i(t_{13} - t_{13}^*) \\ t_{13} + t_{13}^* - i(t_{24} - t_{24}^*) & t_{23} + t_{23}^* - i(t_{14} - t_{14}^*) & t_{11} - t_{22} + t_{33} - t_{44} & t_{34} + t_{34}^* + i(t_{12} - t_{12}^*) \\ t_{14} + t_{14}^* + i(t_{23} - t_{23}^*) & t_{24} + t_{24}^* + i(t_{13} - t_{13}^*) & t_{34} + t_{34}^* - i(t_{12} - t_{12}^*) & t_{11} - t_{22} - t_{33} + t_{44} \\ t_{14} + t_{14}^* + i(t_{23} - t_{23}^*) & t_{24} + t_{24}^* + i(t_{13} - t_{13}^*) & t_{34} + t_{34}^* - i(t_{12} - t_{12}^*) & t_{11} - t_{22} - t_{33} + t_{44} \\ \end{bmatrix} \\ [M] = \begin{bmatrix} m_{11} & m_{12} & m_{13} & m_{14} \\ m_{21} & m_{22} & m_{23} & m_{24} \\ m_{41} & m_{42} & m_{43} & m_{11} + m_{22} - m_{33} - m_{44} & m_{23} + m_{32} + i(m_{14} - m_{41}) & m_{24} + m_{42} - i(m_{13} - m_{32}) \\ m_{14} + m_{41} + i(m_{23} - m_{33}) & m_{24} + m_{32} - i(m_{14} - m_{41}) & m_{11} - m_{22} + m_{33} - m_{44} & m_{43} + i(m_{12} - m_{21}) \\ m_{14} + m_{41} + i(m_{23} - m_{32}) & m_{24} + m_{42} + i(m_{13} - m_{31}) & m_{34} + m_{43} - i(m_{12} - m_{21}) & m_{11} - m_{22} - m_{33} - m_{44} & m_{43} + i(m_{12} - m_{21}) \\ \end{bmatrix} \end{aligned}$$

These represent the most general scattering case in the wave or FSA coordinate system. Note that [T] in the sensor or BSA coordinates has the same eigenvalues as in the wave system, but the eigenvectors are related by a unitary matrix

derived in equation (1.188) and shown again in equation (2.50):

$$\underline{e}_{FSA} = \begin{bmatrix} U_4^C \end{bmatrix} \underline{e}_{BSA} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -i \\ 0 & 0 & i & 0 \end{bmatrix} \underline{e}_{BSA}$$
(2.50)

As before, interest often centres on the special case of backscatter in either the wave or sensor coordinates. In the sensor BSA case the fourth row and columns of [*T*] become zero due to reciprocity ($S_{HV} = S_{VH}$). In both cases [*T*] is reduced to a 3 × 3 coherency matrix, although [*M*] is maintained as 4 × 4. The backscatter mappings in the sensor (BSA) coordinates are then given as follows:

$$[T]_{BSA} = \begin{bmatrix} t_{11} & t_{12} & t_{13} & 0 \\ t_{12}^* & t_{23}^* & t_{33} & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

$$\Rightarrow [M]_{BSA} = \frac{1}{2} \begin{bmatrix} t_{11} + t_{22} + t_{33} & t_{12} + t_{12}^* & t_{13} + t_{13}^* & -i(t_{23} - t_{23}^*) \\ t_{12} + t_{12}^* & t_{11} + t_{22} - t_{33} & t_{23} + t_{23}^* & -i(t_{13} - t_{13}^*) \\ t_{13} + t_{13}^* & t_{23} + t_{23}^* & t_{11} - t_{22} + t_{33} & i(t_{12} - t_{12}^*) \\ -i(t_{23} - t_{23}^*) & -i(t_{13} - t_{13}^*) & i(t_{12} - t_{12}^*) & -t_{11} + t_{22} + t_{33} \end{bmatrix}$$

$$[M]_{BSA} = \begin{bmatrix} m_{11} & m_{12} & m_{13} & m_{14} \\ m_{12} & m_{22} & m_{23} & m_{24} \\ m_{13} & m_{23} & m_{33} & m_{34} \\ m_{14} & m_{24} & m_{34} & m_{44} \end{bmatrix}$$

$$\Rightarrow [T]_{BSA} = \frac{1}{2} \begin{bmatrix} m_{11} + m_{22} + m_{33} - m_{44} & m_{12} - im_{34} & m_{13} + im_{24} \\ m_{12} + im_{34} & m_{11} + m_{22} - m_{33} + m_{44} & m_{23} + im_{14} \\ m_{13} - im_{24} & m_{23} - im_{14} & m_{11} - m_{22} + m_{33} - m_{44} \end{bmatrix}$$

$$(2.52)$$

In the wave (FSA) coordinates we have a different backscatter result, because in this case $S_{\text{HV}} = -S_{\text{VH}}$ and so the third row and column of [*T*] are zero. In this coordinate system the mappings have the following detailed form:

$$[T]_{FSA} = \begin{bmatrix} t_{11} & t_{12} & 0 & t_{14} \\ t_{12}^* & t_{22} & 0 & t_{24} \\ 0 & 0 & 0 & 0 \\ t_{14}^* & t_{24}^* & 0 & t_{44} \end{bmatrix}$$

$$\Rightarrow [M]_{FSA} = \frac{1}{2} \begin{bmatrix} t_{11} + t_{22} + t_{44} & t_{12} + t_{12}^* & i(t_{24} - t_{24}^*) & t_{14} + t_{14}^* \\ t_{12} + t_{12}^* & t_{11} + t_{22} - t_{44} & i(t_{14} - t_{14}^*) & t_{24} + t_{24}^* \\ -i(t_{24} - t_{24}^*) & -i(t_{14} - t_{14}^*) & t_{11} - t_{22} - t_{44} & i(t_{12} - t_{12}^*) \\ t_{14} + t_{14}^* & t_{24} + t_{24}^* & -i(t_{12} - t_{12}^*) & t_{11} - t_{22} + t_{44} \end{bmatrix}$$

$$(2.53)$$

$$[M]_{FSA} = \begin{bmatrix} m_{11} & m_{12} & m_{13} & m_{14} \\ m_{12} & m_{22} & m_{23} & m_{24} \\ -m_{13} & -m_{23} & m_{33} & m_{34} \\ m_{14} & m_{24} & -m_{34} & m_{44} \end{bmatrix}$$

$$\Rightarrow [T]_{FSA} = \frac{1}{2} \begin{bmatrix} m_{11} + m_{22} + m_{33} + m_{44} & m_{12} - im_{34} & m_{14} - im_{23} \\ m_{12} + im_{34} & m_{11} + m_{22} - m_{33} - m_{44} & m_{24} - im_{13} \\ m_{14} + im_{23} & m_{24} + im_{13} & m_{11} - m_{22} - m_{33} + m_{44} \end{bmatrix}$$

$$(2.54)$$

There are then four important results derived directly from the eigenvector expansion of [T]:

1. The conservation of zero wave entropy, introduced after equation (2.32), implies that $\langle [T] \rangle$ has only one non-zero eigenvalue, and so we can state the following important result (Cloude, 1989):

if
$$[S] = f^{-1}([M])$$
 exists then $[M] \to [T] \Rightarrow \lambda_1 > \lambda_2 = \lambda_3 = \lambda_4 = 0$
(2.55)

This constitutes an efficient and complete test of a candidate Mueller matrix: to convert it to its corresponding coherency matrix, calculate the eigenvalue spectrum, and see how many non-zero eigenvalues it has (or equivalently calculate the rank of [T]). If the rank is greater than one then the inverse mapping is not unique.

2. As an extension of this idea we can use the non-negativity of the eigenvalues of [T] as a general test for physical Mueller matrices. It follows that a 4×4 real matrix [M] is a true Mueller matrix if, and only if, its corresponding coherency matrix [T] has a non-negative eigenvalue spectrum so that

$$[M] \to [T] \Rightarrow \lambda_i \ge 0, \quad i = 1, 2, 3, 4 \tag{2.56}$$

3. Experimental measurements of the Mueller matrix [M] are often distorted by errors leading to non-physical results (Cloude, 1986; Brosseau, 1990) (negative scattered powers, degree of polarisation greater than unity, and so on). We can use the eigenvalue spectrum of [T] to devise a filtering scheme for small errors that guarantees a matrix estimate with stable physical properties across the whole Poincaré sphere. The idea is to calculate the eigenvalue spectrum and zero any contributions from negative eigenvalue contributions. The filtering algorithm then has the following form, where sgn is the signum function, +1 for positive values, 0 for 0, and −1 for negative:

$$[M_{\exp}] \rightarrow [T_{\exp}] = \sum_{i} \lambda_i \underline{e}_i \underline{e}_i^{*T} \Rightarrow [\hat{T}]$$
$$= \sum_{i} \frac{1}{2} (1 + sgn(\lambda_i)) \lambda_i \underline{e}_i \underline{e}_i^{*T} \rightarrow [\hat{M}] \qquad (2.57)$$

4. Considering the special case of backscatter, we saw that reciprocity forced a rank reduction of [T]. In particular, the reciprocity theorem forces the third (wave) or fourth (sensor) row/column to be zero (see

equations (2.51) and (2.53)). This leads us directly to the trace condition on backscatter Mueller matrices discussed in equation (2.38). As a special case we consider the possibility of designing wave depolarisers for use in backscatter. We may, for example, consider using a very rough surface or multiple scattering cavity with backscatter geometry. We define the following general form for an isotropic depolariser with $0 \le \delta \le 1$:

$$[M] = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \delta & 0 & 0 \\ 0 & 0 & \delta & 0 \\ 0 & 0 & 0 & \delta \end{bmatrix}$$
(2.58)

This represents a generalization of the example considered in equation (2.35) (which is the special case, $\delta = 0$) where the incident Stokes vector is preserved in form but its degree of polarisation is reduced to δ . In the backscatter direction, reciprocity places the following constraints on δ from equation (2.38)):

Reciprocity

$$\Rightarrow \begin{cases} m_{11} - m_{22} - m_{33} - m_{44} = 1 - 3\delta = 0 \text{ (sensor coordinates)} \\ m_{11} - m_{22} + m_{33} - m_{44} = 1 - \delta = 0 \text{ (wave coordinates)} \end{cases}$$
(2.59)

We see that the ideal depolariser ($\delta = 0$) is in fact impossible to realize in backscatter (it would violate reciprocity). Note that in the BSA system we can, however, obtain a 33% isotropic depolariser. (We return to this idea in Section 2.5.) Consequently any depolarisation in backscattering must be anisotropic; that is, it must transform as well as depolarise. For this reason it is interesting to consider a more detailed parameterization of depolarisation, as follows.

2.4 General theory of scattering entropy

In this section we formulate a general model of depolarisation that scales to arbitrary dimension of coherency matrix $N \times N$. The basic idea is to always identify the 'polarising' contribution with the dominant eigenvector of the matrix—the eigenvector corresponding to the largest eigenvalue. The other eigenvectors then contribute to depolarisation with a strength given by the remaining minor eigenvalues. By employing multidimensional unitary transformations we will then be able to parameterize all possible types of system depolarisation behaviour. We start with the general formulation and then specialize it to the three important cases for N = 2, 3 and 4. We then consider the effects of scattering symmetries on constraining the degrees of freedom involved in both polarised and depolarised components.

The starting point for our analysis is the idea of a unitary reduction operator: a matrix $[U_{-1}]$ which acts to reduce the dimensionality of an N × N unitary matrix to $N - 1 \times N - 1$, as shown in equation (2.60) (Cloude, 1986, 1995b, 2001a):

$$\begin{bmatrix} U_{-1} \end{bmatrix} \begin{bmatrix} U_N \end{bmatrix} = \begin{bmatrix} 1 & \underline{0}^T \\ \underline{0} & U_{N-1} \end{bmatrix} = \begin{bmatrix} 1 & \underline{0}^T \\ \underline{0} & \exp(iH_{N-1}) \end{bmatrix}$$
(2.60)

$$H_{N-1} = \sum_{k=1}^{M} h_k \Psi_k \implies \underline{h} = depolarisation \ state \ vector$$

. .

This operator acts to convert the first column of the coherency matrix to the identity. The next stage is to apply such an operator to the matrix of ranked eigenvectors \underline{e}_i of the coherency matrix

$$[U_N] = \begin{bmatrix} \underline{e}_1 & \underline{e}_2 & \dots & \underline{e}_N \end{bmatrix}$$
(2.61)

and to identify the submatrix U_{N-1} with the depolarising aspects of the scattering process. In this way U_{N-1} involves continuous smooth transformation away from the polarised reference state (the dominant eigenvector). The submatrix U_{N-1} may be further parameterized in terms of an N - 1 × N - 1 Hermitian matrix, related to the unitary transformation by a matrix exponential, and itself conveniently expanded in terms of a set of scalar parameters being the basis elements of the underlying algebra. This vector of coefficients we then term the depolarisation state vector (see equation (2.60)), as it determines the nature of the depolarisation, with different state vectors representing different types and degrees of depolarised scattering.

This then leads us to the following notation to characterize the number of parameters involved in polarising and depolarising components of the decomposition of a general N \times N coherency matrix $T_{\rm N}$:

$$T_{\rm N} = ({\rm E} + {\rm L}) + [{\rm E} + {\rm L}]$$
 (2.62)

where [..] are the depolarising parameters and (..) the polarising terms, and for each, E are those parameters associated with the eigenvectors and L the eigenvalues. From the general structure of $N \times N$ coherency matrices we then have the following constraints:

- [L] + (L) = N
- $[E] + (E) = \dim(SU(N)) \operatorname{rank}(SU(N)) = N(N 1)$

Note that the total number of eigenvector parameters is given by the sum [E] $+ (E) = \dim(SU(N)) - r(SU(N))$, where $\dim() = N^2 - 1$ is the dimension of the group and r = N - 1 is the rank of the Cartan sub-algebra or the number of mutually commuting generators (see Appendix 2). For example, for N = 1 there are no useful eigenvector parameters, and for N = 2 we only have two, while for N = 4 (the most general bistatic scattering case) we have twelve parameters available. We now look at special cases to illustrate how this classification scheme works, and in particular how general scattering symmetries restrict the number of parameters.

2.4.1 Depolarisation for N = 2 scattering systems

In the simplest non-trivial case of N = 2, each eigenvector can be parameterized in the form shown in equation (2.63):

$$\underline{e}_1 = e^{i\chi} \begin{bmatrix} \cos \alpha \\ \sin \alpha e^{i\phi_1} \end{bmatrix}$$
(2.63)

The unitary reduction operator can be constructed from these parameters as shown in equation (2.64):

$$e^{-i\chi} \begin{bmatrix} \cos\alpha & \sin\alpha \\ -\sin\alpha & \cos\alpha \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & e^{-i\phi_1} \end{bmatrix} \begin{bmatrix} \underline{e}_1 & \underline{e}_2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & e^{-i\phi_1} \end{bmatrix}$$
(2.64)

Here we see that the depolarisation subspace is now one-dimensional and consists of a scalar phase shift. This, however, remains a 'hidden' parameter (derived from the Cartan sub-algebra), since phase shifts of the original polarised state are not observed when forming the coherency matrix. This leads to the following decomposition of general two-dimensional coherency matrices (see also equation (2.12)):

$$\mathbf{T}_2 = (2+1) + [0+1] \tag{2.65}$$

Note that there are no depolarisation parameters in the eigenvectors, and that all depolarisation effects lie in the eigenvalue spectrum. This leads to the wave dichotomy and degree of polarisation ideas discussed in Section 2.1.1. More interesting possibilities exist for higher-dimensional matrices, as we now consider.

2.4.2 Depolarisation for N = 3 scattering systems

For the case N = 3 we can parameterize the general eigenvector as a natural extension of the N = 2 case, as shown in equation (2.66):

$$\underline{e}_{1} = e^{i\chi} \begin{bmatrix} \cos \alpha \\ \sin \alpha \cos \psi e^{i\phi_{1}} \\ \sin \alpha \sin \psi e^{i\phi_{2}} \end{bmatrix}$$
(2.66)

This can be used to parameterize backscatter problems, when the reciprocity theorem forces the coherency matrix to be 3×3 (see equations (2.51) and (2.52)). From this we can define the unitary reduction operator as a cascade of three matrices shown in equation (2.67):

$$e^{-i\chi} \begin{bmatrix} \cos\alpha & \sin\alpha & 0\\ -\sin\alpha & \cos\alpha & 0\\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0\\ 0 & \cos\psi & \sin\psi\\ 0 & -\sin\psi & \cos\psi \end{bmatrix} \begin{bmatrix} 1 & 0 & 0\\ 0 & e^{-i\phi_1} & 0\\ 0 & 0 & e^{-i\phi_2} \end{bmatrix}$$
$$\times \begin{bmatrix} \underline{e}_1 & \underline{e}_2 & \underline{e}_3 \end{bmatrix} = \begin{bmatrix} 1 & \underline{0}^T\\ \underline{0} & U_2 \end{bmatrix}$$
(2.67)

In this case [T]/[M] have a maximum of nine parameters, and SU(3) is the governing unitary group for the eigenvectors of [T]. However, the unitary reduction operator means that all depolarisation effects are controlled by SU(2), which has dimension 3 and rank 1. Hence in N = 3 backscatter the polarised/depolarised decomposition can be written in compact form as follows:

$$\mathbf{T}_3 = (4+1) + [2+2] \tag{2.68}$$

This shows that the eigenvectors now contain two depolarising parameters, complemented by two real eigenvalues. The polarised component has four eigenvector parameters and one eigenvalue (its amplitude). Note that this decomposition makes no assumptions about scattering symmetry (other than reciprocity). We now turn to consider the restrictions placed by scattering symmetries on the form of the coherency and Mueller matrices in backscatter.

2.4.2.1 Backscatter symmetries and depolarisation

One of the most important symmetries in scattering theory is reciprocity. We have seen that the effect of this symmetry in backscatter problems is to generate a rank reduction of the coherency matrix. In this section we consider the form of [T] for various other scattering symmetry assumptions in the medium, with a view to investigating any further simplifications that may occur in the form of [T]. There are three primary types of symmetry to be considered (Perrin, 1942; Van de Hulst, 1981; Nghiem, 1992; Cloude, 1996), as illustrated in Figure 2.4. Here we see a representation of the plane of polarisation of the incident wave and consider the scattering to be due to the sum of (independent) contributions from a large number of scatterers in the scene. The first symmetry we can consider is reflection symmetry. Here we assume that for every scatterer at point P there is a matching scatterer in the reflected position about some axis AA'. Note that AA' need not be aligned with the sensor H and V coordinates in which case it is parameterized by a rotation angle θ .

Due to the independence assumption, the coherency matrix for this scene is given by the sum of coherency matrices from P and Q. Assuming for the moment that AA' is aligned ($\theta = 0$), then, as shown in equation (1.145), the <u>k</u> vectors at P and Q are related by a minus sign in their crosspolarised components, so that the coherency matrix has zero elements in some off-diagonal positions, as shown in equation (2.69).

$$\underline{k}_{P} = \begin{bmatrix} k_{0} \\ k_{1} \\ k_{2} \end{bmatrix} \Rightarrow \underline{k}_{Q} = \begin{bmatrix} k_{0} \\ k_{1} \\ -k_{2} \end{bmatrix}$$
$$\Rightarrow [T] = [T_{P}] + [T_{Q}] = \begin{bmatrix} t_{11} & t_{12} & t_{13} \\ t_{12}^{*} & t_{22} & t_{23} \\ t_{13}^{*} & t_{23}^{*} & t_{33} \end{bmatrix} + \begin{bmatrix} t_{11} & t_{12} & -t_{13} \\ t_{12}^{*} & t_{22} & -t_{23} \\ -t_{13}^{*} & -t_{23}^{*} & t_{33} \end{bmatrix}$$
$$= \begin{bmatrix} t_{11} & t_{12} & 0 \\ t_{12}^{*} & t_{22} & 0 \\ 0 & 0 & t_{33} \end{bmatrix}$$
(2.69)



Fig. 2.4 Reflection and rotation symmetry

In this case we see that one consequence of reflection symmetry is that the coherence between co- and crosspolarised channels will be zero: $\langle S_{HH} S_{HV}^* \rangle = \langle S_{VV} S_{VH}^* \rangle = 0$, and so on. Reintroducing the idea that θ may not be zero produces the following general form for a reflection symmetric medium:

$$[T_{\theta}] = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos 2\theta & -\sin 2\theta \\ 0 & \sin 2\theta & \cos 2\theta \end{bmatrix} \begin{bmatrix} t_{11} & t_{12} & 0 \\ t_{12}^* & t_{22} & 0 \\ 0 & 0 & t_{33} \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos 2\theta & \sin 2\theta \\ 0 & -\sin 2\theta & \cos 2\theta \end{bmatrix}$$
(2.70)

Note that this will introduce apparent correlations between co- and crosschannels due to crosspolarisation (but not depolarisation) by the axis misalignment.

The second type of symmetry to consider is rotation symmetry. Here we consider the form of [T] that remains invariant under arbitrary rotations in the plane of polarisation. By definition, and as an extension of the discussion around equation (2.69), [T] must then be formed from linear combinations of the outer product of eigenvectors of the plane rotation matrix (see Section 4.1.1). In this case we can therefore predict the general form of a rotation symmetric [T] as shown in equation (2.71):

$$[T] = a\underline{\nu}_1\underline{\nu}_1^{*T} + b\underline{\nu}_2\underline{\nu}_2^{*T} + c\underline{\nu}_3\underline{\nu}_3^{*T} = \begin{bmatrix} a & 0 & 0\\ 0 & b+c & i(b-c)\\ 0 & -i(b-c) & b+c \end{bmatrix}$$
(2.71)

Finally we consider the most general symmetry, formed by insisting that every axis AA', for all θ , has the PQ reflection symmetry. This is termed azimuthal symmetry, and the form of [*T*] can be found by adding reflection symmetric versions of the rotation matrix of equation (2.71) to obtain equation (2.72):

$$[T] = \begin{bmatrix} a & 0 & 0 \\ 0 & b+c & i(b-c) \\ 0 & -i(b-c) & b+c \end{bmatrix} + \begin{bmatrix} a & 0 & 0 \\ 0 & b+c & -i(b-c) \\ 0 & i(b-c) & b+c \end{bmatrix}$$
$$= 2\begin{bmatrix} a & 0 & 0 \\ 0 & b+c & 0 \\ 0 & 0 & b+c \end{bmatrix}$$
(2.72)

Here we have obtained a very simple diagonal coherency matrix, with two equal diagonal values. We now turn to consider the polarised/depolarised decomposition for each of these symmetry cases.

2.4.2.2 Depolarisation under azimuthal backscattering symmetry

This severe symmetry assumption leads directly to a diagonal coherency (and Mueller) matrix with two degenerate coherency eigenvalues. Consequently, [T]/[M] have only two free parameters, and the coherency/covariance and (BSA) Mueller matrices for backscatter can be written as shown in
equation (2.73):

$$[T_{3}] = \begin{bmatrix} m & 0 & 0 \\ 0 & m\kappa & 0 \\ 0 & 0 & m\kappa \end{bmatrix}$$

$$\Leftrightarrow [C_{3}] = \frac{m}{2} \begin{bmatrix} 1+\kappa & 0 & 1-\kappa \\ 0 & 2\kappa & 0 \\ 1-\kappa & 0 & 1+\kappa \end{bmatrix}$$

$$\Leftrightarrow [M] = \frac{m}{2} \begin{bmatrix} 1+2\kappa & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 2\kappa - 1 \end{bmatrix}$$

$$(2.73)$$

In this symmetry assumption the polarising/depolarising decomposition therefore reduces to equation (2.74):

$$T_{3az} = (0+1) + [0+1]$$
(2.74)

In this case only two parameters are required to characterize the medium (such as VV power and HH/VV coherence *or* HV, but *not* both HV and HH/VV coherence, as these two are now related). A convenient parameterization for the most usual situation in remote sensing applications (a > b + c in equation (2.72)) is provided in terms of the polarised power (*m*) and depolarisation or noise parameter (κ) shown in equation (2.73).

Note that there are two important special cases. When $\kappa = 0$ we have zero depolarisation and backscatter from a sphere symmetric scatterer (or random collection of such), and at the other extreme $\kappa = 1$, and [*T*] reduces to the 3×3 identity matrix. This represents the most depolarising case. Note, however, that the Mueller matrix for this case still does not equal a noise source ideal depolariser (equation (2.35)); in fact, it corresponds to a 33% isotropic depolariser. As mentioned in the discussion around equation (2.59), this is due to the fact that HV and VH remain correlated for all backscatter problems, and hence an identity coherency matrix corresponds not to equal noise power in all channels, but with the cross-channel having only half the power of the copolarised channels. This indicates that a proper analysis of polarised signals in the presence of noise requires a full 4×4 coherency matrix analysis. This residual structure implies that even in the most depolarising backscatter case there will be some residual information on the scatterer properties, even just that following from the reciprocity theorem.

2.4.2.3 Depolarisation under reflection symmetry

In this more complicated case, both [T] and [M] have six free parameters, and the coherency matrix can be written in the general form shown in equation (2.75), where θ is the angle mismatch between radar coordinates and the axis of symmetry (for example the local normal in surface scattering; see Section 3.1):

$$[T_{\theta}] = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos 2\theta & -\sin 2\theta \\ 0 & \sin 2\theta & \cos 2\theta \end{bmatrix} \begin{bmatrix} t_{11} & t_{12} & 0 \\ t_{12}^* & t_{22} & 0 \\ 0 & 0 & t_{33} \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos 2\theta & \sin 2\theta \\ 0 & -\sin 2\theta & \cos 2\theta \end{bmatrix}$$
(2.75)

The polarised/depolarised decomposition can be written for reflection symmetry as shown in equation (2.76):

$$[T]_{\rm ref} = (3+1) + [0+2] \tag{2.76}$$

This result shows formally that for backscatter from random media with reflection symmetry [E] = 0 and [L] = 2, so the eigenvectors contain *no* information about depolarisation, only about the polarising influence of the dominant eigenvector. This results because the 3×3 matrix of eigenvectors for such media can be written as a function of a 2×2 unitary matrix, as shown in equation (2.77):

$$\begin{bmatrix} U_3 \end{bmatrix} = \begin{bmatrix} \begin{bmatrix} U_2 \end{bmatrix} & 0\\ 0 & 1 \end{bmatrix} \Rightarrow \begin{bmatrix} U_{-1} \end{bmatrix} \begin{bmatrix} U_3 \end{bmatrix} = \begin{bmatrix} 1 & 0\\ 0 & 1 \end{bmatrix}$$
(2.77)

All the depolarisation information for such media is therefore contained in the two minor eigenvalues of [*T*]. These eigenvalues are not usually used themselves as measures of depolarisation in such media. Instead, widespread use is made of two functions of the eigenvalues λ_i : scattering entropy and anisotropy (Pottier, 1997), defined as in equations (2.78) and (2.79):

$$H = -\sum_{i=1}^{3} P_i \log_3 P_i$$
 $P_i = \frac{\lambda_i}{\sum \lambda}$ $0 \le H \le 1$ (2.78)

$$A = \frac{\lambda_2 - \lambda_3}{\lambda_2 + \lambda_3} \quad 0 \le A \le 1 \tag{2.79}$$

The entropy is zero for zero depolarisation; that is, when [T] has only one non-zero eigenvalue. At the other extreme, H = 1 when [T] is diagonal and maximum depolarisation occurs. The anisotropy is a useful parameter for assessing the type of symmetry. For azimuthal symmetry the minor eigenvalues are equal, and so A = 0. On the other hand, if only the smallest eigenvalue falls to zero then A = 1, and we have significant departure from azimuthal symmetry. This situation occurs, for example, for in-plane scattering from one-dimensional rough surfaces and in scattering by chiral particles (see Chapter 3).

2.4.2.4 The entropy/alpha decomposition for N = 3

Despite the completeness of the full unitary reduction operator approach, in all but the azimuthal symmetry case it requires estimation of multiple parameters from the data. Very often a simpler approach to the parameterization of depolarisation effects is all that is required. The basic motivation here is to seek a method that provides just two parameters—one indicating the total level of depolarisation, and the other the average polarised information. In addition, this method must cope not only with azimuthal and reflection symmetry, but also the most general case. It is desirable that the parameters be as robust as possible, so we require them to be polarisation basis invariant, otherwise we would need to change the interpretation every time we changed the reference polarisation base. This again suggests an eigenvalue decomposition of $\langle [T] \rangle$, as shown in equation (2.80), but this time, in contrast to the unitary reduction operator approach, we seek to find averages that may be used for a reduced parameterization (Cloude, 1995c, 1996, 1997a; Lee, 2008).

$$\langle [T] \rangle = \begin{bmatrix} \underline{e}_1 & \underline{e}_2 & \underline{e}_3 \end{bmatrix} \begin{bmatrix} \lambda_1 & 0 & 0\\ 0 & \lambda_2 & 0\\ 0 & 0 & \lambda_3 \end{bmatrix} \begin{bmatrix} \underline{e}_1 & \underline{e}_2 & \underline{e}_3 \end{bmatrix}^{*T}$$
(2.80)

The spread of total scattered power across the eigenvalues is evidently a good indicator of depolarisation. This can be turned into a quantitative measure by first normalizing the eigenvalues to unit sum, so we define three parameter P_i , as shown in equation (2.81). These can then be interpreted as probabilities of the statistically independent 'polarised' states, given by the eigenvectors \underline{e}_i . The spread of probabilities can then be represented by a single scalar, the entropy being defined as shown in equation (2.81):

$$0 \le P_i = \frac{\lambda_i}{\sum \lambda} \le 1 \Rightarrow H = -\sum_{i=1}^3 P_i \log_3 P_i \quad 0 \le H \le 1$$
(2.81)

This then provides a suitable *depolarising* parameter; which is zero for zero depolarisation (while still allowing crosspolarisation), and one for maximum depolarisation. But what of the polarised component? From each eigenvector we can select the scattering mechanism α as a suitable basic invariant polarised parameter, as shown in equation (2.82).

$$\underline{e} = \begin{bmatrix} e_1 \\ e_2 \\ e_3 \end{bmatrix} = e^{i\chi} \begin{bmatrix} \cos \alpha \\ \sin \alpha \cos \psi e^{i\phi_1} \\ \sin \alpha \sin \psi e^{i\phi_2} \end{bmatrix} \Rightarrow \alpha = \cos^{-1}(|e_1|) \quad 0 \le \alpha \le 90^{\circ}$$
(2.82)

However, we have three such values—one for each eigenvector, as shown in equation (2.83):

$$[U_3] = \begin{bmatrix} \cos\alpha_1 & \cos\alpha_2 & \cos\alpha_3\\ \sin\alpha_1\cos\psi_1e^{i\delta_1} & \sin\alpha_2\cos\psi_2e^{i\delta_2} & \sin\alpha_3\cos\psi_3e^{i\delta_3}\\ \sin\alpha_1\sin\psi_1e^{i\gamma_1} & \sin\alpha_2\sin\psi_2e^{i\gamma_2} & \sin\alpha_3\sin\psi_3e^{i\gamma_3} \end{bmatrix}$$
(2.83)

The statistical interpretation suggests forming an average as a sum of the three values, weighted by their probabilities, so we can form the average alpha parameter as shown in equation (2.84):

$$\overline{\alpha} = P_1 \alpha_1 + P_2 \alpha_2 + P_3 \alpha_3 \tag{2.84}$$

Note that the form of the unitary matrix of eigenvectors shown in equation (2.83) is overparameterized. The unitary reduction operator provides a much more rigorous approach (for example, the columns of the unitary matrix are not independent, as implied in equation (2.83), as they are all mutually orthogonal). However, the approach in equation (2.83) has a strong physical appeal in terms of the probabilistic interpretation, and provides a useful parameterization of the mean scattering mechanism. We have therefore derived a suitable pair of parameters: entropy H and scattering mechanism $\overline{\alpha}$. To aid interpretation of different types of polariser/depolariser, use can then be made of the two-dimensional H/ α or entropy/alpha plane, as shown in Figure 2.5 (Cloude, 1996).



Fig. 2.5 Entropy/alpha diagram for N = 3 backscatter

All reciprocal backscattering problems can be represented as points in this plane. However, not all of the plane represents physical depolarisers. Because of the averaging inherent in equation (2.84) there are bounds on the maximum and minimum alpha angle we can obtain for a given entropy H. Only for zero entropy can we obtain the full 90 degrees of alpha variation. For other values of entropy we can calculate the maximum and minimum values of alpha to define two bounding curves, shown as I and II in Figure 2.5. The lower bounding curve is defined by azimuthally symmetric scattering. These have $\alpha = 0$ for the dominant eigenvector, and equal 'spill-over' into the $\alpha = \pi/2$ eigenvectors by the depolarising parameter 'm' as shown in equation (2.85). The upper bound can be found in a similar manner by first 'switching off' the signal in the $\alpha = 0$ eigenvector, and then only when the $\alpha = \pi/2$ subspace is full, allowing signal energy into this eigenvector. The resulting bounding curve II is also shown in equation (2.85). Both curves tend to the same alpha point for maximum entropy H = 1, namely $\pi/3$:

$$CURVE I [T]_{I} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & m & 0 \\ 0 & 0 & m \end{bmatrix} \quad 0 \le m \le 1$$

$$CURVE II \begin{cases} [T]_{II} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2m \end{bmatrix} \le m \le 0.5 \quad (2.85)$$

$$[T]_{II} = \begin{bmatrix} 2m - 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad 0.5 \le m \le 1$$

Although the main motivation behind the H/ α decomposition is to derive a small set of key parameters to simplify classification of depolarisers, the basic scheme is often augmented by the addition of a third parameter: the scattering anisotropy A (equation (2.79))—again a basis invariant depolarising parameter,

as shown in equation (2.86).

$$A = \frac{\lambda_2 - \lambda_3}{\lambda_2 + \lambda_3} \Rightarrow \begin{cases} a_1 = (1 - H)(1 - A) \\ a_2 = (1 - H)A \\ a_3 = HA \\ a_4 = H(1 - A) \end{cases}$$
(2.86)

With entropy and anisotropy now defined, several combinations of these two have been proposed as suitable for the enhancement of important special cases of depolarisation. These are shown as a_1 to a_4 in equation (2.86). These parameters have large values for the following cases:

- a_1 This parameter is large for single dominant eigenvalue scattering (H small) with added noise (A small).
- a_2 This parameter is large for a strong single polarising element (H small) in the presence of a second weaker mechanism (A large).
- *a*₃ This parameter is large for two strong mechanisms in the presence of a weak third.
- *a*⁴ This parameter is large for strongly depolarising systems with three roughly equal eigenvalues.

This entropy/alpha approach to depolarisation was first developed for radar backscatter, where it is often termed the Cloude–Pottier decomposition, after the original authors (Cloude, 1996, 1997a). One advantage of this entropy/alpha approach is that it can be scaled to different coherency dimensions, as we now discuss for the case of reduced dimension N = 2.

2.4.2.5 The entropy/alpha decomposition for N = 2

While the entropy/alpha approach was originally designed to simplify multiparameter depolarisation, as occurs in N = 3 and N = 4 problems, it can also be applied to the simpler case of N = 2 depolarisation (Ainsworth, 2007; Cloude, 2007b). Using the interpretation of normalized eigenvalues as probabilities P_i , together with the fact that in N = 2 problems the second eigenvector can be derived entirely from the principal vector using orthogonality, we obtain an entropy/alpha parameterization of the wave coherency matrix [J], as shown in equation (2.87):

$$[J] = \begin{bmatrix} J_{xx} & J_{xy} \\ J_{xy}^* & J_{yy} \end{bmatrix} \Rightarrow \begin{cases} [U_2] = \begin{bmatrix} \cos \alpha & -\sin \alpha e^{-i\delta} \\ \sin \alpha e^{i\delta} & \cos \alpha \end{bmatrix} \\ [D] = (\lambda_1 + \lambda_2) \begin{bmatrix} P_1 & 0 \\ 0 & P_2 \end{bmatrix} \\ \Rightarrow \begin{cases} \overline{\alpha_2} = P_1 \alpha + P_2 \left(\frac{\pi}{2} - \alpha\right) = \alpha (P_1 - P_2) + P_2 \frac{\pi}{2} \\ H_2 = \sum_{i=1}^2 P_i \log_2 P_i \end{cases}$$

$$(2.87)$$

Again we can define an entropy/alpha plane as shown in Figure 2.6. In this case the upper and lower bounds can be found by considering the parameterized



Fig. 2.6 Entropy/alpha diagram for N = 2 depolarisation

diagonal 2×2 matrix shown in equation (2.88):

$$\begin{bmatrix} 1 & 0 \\ 0 & m \end{bmatrix} \quad 0 \le m \le 1 \Rightarrow \overline{\alpha}_{\min} = \frac{m\pi}{2(1+m)}$$

$$\begin{bmatrix} m & 0 \\ 0 & 1 \end{bmatrix} \quad 0 \le m \le 1 \Rightarrow \overline{\alpha}_{\max} = \frac{\pi}{2(1+m)}$$
(2.88)

In this case the lower and upper curves show a symmetry about $\alpha = 45$ degrees. This can be traced to the two-dimensional eigenvector space, one of which has $\alpha = 0$ and the other $\pi/2$. In the N = 3 case we saw a 2:1 bias in favour of the $\pi/2$ eigenvector subspace, which tends to lift the alpha range with increasing entropy (Figure 2.5).

In this N = 2 case we also have a simple geometrical interpretation of α on the Poincaré sphere (see Figure 1.12). In the special case that x and y in equation (2.87) represent polarisation wave states (see Section 1.33), the angles α and δ are then related to the orientation and ellipticity angles θ , τ of the polarisation ellipse via a spherical triangle construction on the Poincaré sphere. However, here we have generalized this idea so that x and y can now be arbitrary channels of data; that is, two elements of the [S] matrix, anticipating the ideas of compact polarimetry (discussed more fully in Section 9.3.4).

The averaging implied by the entropy/alpha approach does not pick out the state P corresponding to the maximum eigenvalue as in the polarised/depolarised decomposition, but instead forms an average based on a probabilistic interpretation of making measurements on the wave and obtaining polarisation X and Y with probabilities P_1 and P_2 respectively. Hence the average polarisation state would have a corresponding alpha value given by $\overline{\alpha}$. For example, when the coherency matrix approaches the identity (noise), then $\overline{\alpha} = \frac{\pi}{4}$, being a mixture of the state X ($\alpha = 0$) with its antipodal orthogonal state Y ($\alpha = \pi/2$).

In radar there are two important special cases when the N = 2 formalism becomes important. It is often advantageous in radar design (from a cost, data rate and coverage point of view) to employ a single transmitted polarisation state and a coherent dual channel receiver to measure orthogonal components of the scattered signal (see Chapter 9) (Souyris 2005; Raney, 2006, 2007). Such dual polarised radars are not capable of reconstructing the complete scattering matrix, but instead can be used to reconstruct a column of the [S] matrix or more generally some projection. From this we can then construct a 2×2 coherency matrix [J] as an example of N = 2 depolarisation. One key decision in the design of such radars is the best single polarisation to employ as the reference point X on the Poincaré sphere. For example, it is shown in Chapter 4 that circular polarisation would be a good choice, since the coherence between coand cross-circular polarisation can be used to estimate orientation of the scatterer. However, circularly polarised transmitters are not widely employed in radar imaging systems, where there is a preference for linear polarisations. For example, several radar systems employ horizontal (H) or vertical (V) polarisation transmit and receive H and V components. These radars can be used to estimate the following forms of the coherency matrix:

$$[J_H] = \begin{bmatrix} \langle S_{HH} S_{HH}^* \rangle & \langle S_{HH} S_{HV}^* \rangle \\ \langle S_{HV} S_{HH}^* \rangle & \langle S_{HV} S_{HV}^* \rangle \end{bmatrix} \quad [J_V] = \begin{bmatrix} \langle S_{VV} S_{VV}^* \rangle & \langle S_{VV} S_{VH}^* \rangle \\ \langle S_{VH} S_{VV}^* \rangle & \langle S_{VH} S_{VH}^* \rangle \end{bmatrix}$$
(2.89)

Note that these represent important examples of the generalization of the wave coherency matrix. We can summarize the polarimetric information content of these new 2×2 matrices by relating them to the full coherency matrix [*T*], as shown in equation (2.90):

_

- 7

$$\underline{k}_{H} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \underline{k} \Rightarrow [J_{H}] = \langle \underline{k}_{H} \underline{k}_{H}^{*T} \rangle = \frac{1}{2} \begin{bmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} [T] \begin{bmatrix} 1 & 0 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}$$
$$= \frac{1}{2} \begin{bmatrix} t_{11} + t_{22} + 2\operatorname{Re}(t_{12}) & t_{13} + t_{23} \\ (t_{13} + t_{23})^{*} & t_{33} \end{bmatrix}$$
$$\underline{k}_{V} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \underline{k} \Rightarrow [J_{V}] = \langle \underline{k}_{V} \underline{k}_{V}^{*T} \rangle = \frac{1}{2} \begin{bmatrix} 1 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix} [T] \begin{bmatrix} 1 & 0 \\ -1 & 0 \\ 0 & 1 \end{bmatrix}$$
$$= \frac{1}{2} \begin{bmatrix} t_{11} + t_{22} - 2\operatorname{Re}(t_{12}) & t_{13} - t_{23} \\ (t_{13} - t_{23})^{*} & t_{33} \end{bmatrix}$$
(2.90)

To emphasize the importance of the off-diagonal elements of these 2×2 matrices, consider the effects of rotation about the line of sight, as occurs, for example, in scattering from a sloped surface (see Chapter 3). To see this we consider the form of the dual polarisation scattering vectors for coherent point scatterers in terms of a rotation about the line of sight. These are shown in equation (2.91) as projections of the coherent Pauli *k* vectors:

$$\underline{k}_{H} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos 2\theta & -\sin 2\theta \\ 0 & \sin 2\theta & \cos 2\theta \end{bmatrix} \cdot \begin{bmatrix} k_{0} \\ k_{1} \\ 0 \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} k_{0} + k_{1} \cos 2\theta \\ k_{1} \sin 2\theta \end{bmatrix}$$
$$\underline{k}_{V} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos 2\theta & -\sin 2\theta \\ 0 & \sin 2\theta & \cos 2\theta \end{bmatrix} \cdot \begin{bmatrix} k_{0} \\ k_{1} \\ 0 \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} k_{0} - k_{1} \cos 2\theta \\ k_{1} \sin 2\theta \end{bmatrix}$$
(2.91)

It follows that the coherency matrices for reflection symmetric random media take the form shown in equation (2.92), from which we note the presence of complex off-diagonal terms.

$$\begin{split} [J_{H}] &= \left\langle \underline{k}_{H} \underline{k}_{H}^{*T} \right\rangle \\ &= \frac{1}{2} \left\langle \begin{bmatrix} |k_{0}|^{2} + |k_{1}|^{2} \cos^{2} 2\theta + 2 \operatorname{Re}(k_{0}k_{1}^{*}) \cos 2\theta & |k_{1}|^{2} \cos 2\theta \sin 2\theta + k_{0}k_{1}^{*} \sin 2\theta \\ |k_{1}|^{2} \cos 2\theta \sin 2\theta + k_{1}k_{0}^{*} \sin 2\theta & |k_{1}|^{2} \sin^{2} 2\theta \end{bmatrix} \right\rangle \\ [J_{V}] &= \left\langle \underline{k}_{V} \underline{k}_{V}^{*T} \right\rangle \\ &= \frac{1}{2} \left\langle \begin{bmatrix} |k_{0}|^{2} + |k_{1}|^{2} \cos^{2} 2\theta - 2 \operatorname{Re}(k_{0}k_{1}^{*}) \cos 2\theta & |k_{1}|^{2} \cos 2\theta \sin 2\theta - k_{0}k_{1}^{*} \sin 2\theta \\ |k_{1}|^{2} \cos 2\theta \sin 2\theta - k_{1}k_{0}^{*} \sin 2\theta & |k_{1}|^{2} \sin^{2} 2\theta \end{bmatrix} \right\rangle \\ (2.92) \end{split}$$

Only for zero tilt ($\theta = 0$) or uniformly random θ do these diagonal terms disappear. This suggests that the phase terms will be important for identifying coherent point scatterers. Using a coherent (zero entropy) assumption we can also relate the dual polarisation alpha parameter to the scattering matrix elements for a symmetric point ($S_{XY} = 0$; see equation (4.12)) as shown in equation (2.93):

$$\tan \alpha_2 = \left| \frac{(S_{XX} - S_{YY}) \sin 2\theta}{(S_{XX} + S_{YY}) + (S_{XX} - S_{YY}) \cos 2\theta} \right|$$

$$\delta_2 = \arg \left(\frac{(S_{XX} - S_{YY}) \sin 2\theta}{(S_{XX} + S_{YY}) + (S_{XX} - S_{YY}) \cos 2\theta} \right)$$
(2.93)

In the more general case of a non-symmetric point scatterer these relations take the form shown in equation (2.94):

$$\tan \alpha_{2} = \left| \frac{2S_{XY} \cos 2\theta + (S_{XX} - S_{YY}) \sin 2\theta}{(S_{XX} + S_{YY}) + (S_{XX} - S_{YY}) \cos 2\theta - 2S_{XY} \sin 2\theta} \right|$$

$$\delta_{2} = \arg \left(\frac{2S_{XY} \cos 2\theta + (S_{XX} - S_{YY}) \sin 2\theta}{(S_{XX} + S_{YY}) + (S_{XX} - S_{YY}) \cos 2\theta - 2S_{XY} \sin 2\theta} \right)$$
(2.94)

One advantage of the entropy/alpha approach is how it scales naturally with changing dimension. We have shown above how the original N = 3 idea can be modified for the restricted N = 2 case, and so now finally turn to consider application of the concept to the most general bistatic N = 4 scattering case. First, however, we consider the general properties of bistatic depolarisation.

2.4.3 Depolarisation in N = 4 scattering systems

In the general bistatic scattering case, [T]/[M] have up to sixteen parameters and SU(4) is the governing unitary group. SU(4) has dimension 16 and rank 4, so [E] + (E) = 16 - 4 = 12, and [L] + (L) = 4. By application of the unitary reduction operator, depolarisation in bistatic scattering systems is then controlled by [L] = 3 eigenvalues and the SU(3) group for eigenvectors. SU(3) has dimension 8 and rank 2, so that we can write the polarising/depolarising decomposition in compact form as shown in equation (2.95), which shows that there are now up to six eigenvector parameters associated with depolarisation.

$$[T]_{\text{bistatic}} = (6+1) + [6+3] \tag{2.95}$$

These can be generated from the Gell–Mann matrices (see Appendix 2) (Cloude, 1986, 1995b; Ferro–Famil, 2000). As in the N = 2 case, however, there are several important symmetries that further reduce the number of parameters in bistatic scattering.

2.4.3.1 Bistatic scattering symmetries

The first symmetry to consider is the reciprocity theorem. In backscatter problems this led to a symmetric BSA scattering matrix and rank-3 coherency matrix. In bistatic problems its effect is more subtle. We begin by considering the effect of the theorem on the [S] matrix itself (Saxon, 1955; Mishchenko, 2000). In equation (2.96) we show that the effect of interchange of transmitter and receiver involves a change of coordinates plus a transpose operation:

$$\frac{\underline{E}_{S}^{A}}{\underline{E}_{S}^{B}} = \begin{bmatrix} S \left(\underline{\beta}_{i} \quad \underline{\beta}_{s} \right) \end{bmatrix} \cdot \underline{P}_{A} \\
\underline{E}_{S}^{B} = \begin{bmatrix} S \left(-\underline{\beta}_{i} \quad -\underline{\beta}_{s} \right) \end{bmatrix} \cdot \underline{P}_{B} \\
Reciprocity Theorem \\
\Rightarrow \begin{bmatrix} S \left(\underline{\beta}_{i} \quad \underline{\beta}_{s} \right) \end{bmatrix} = \begin{bmatrix} S \left(-\underline{\beta}_{i} \quad -\underline{\beta}_{s} \right) \end{bmatrix}^{T} \quad (2.96)$$

In equation (2.97) we show explicitly how the elements of the [S] matrices are related:

$$[S] = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} S \left(-\underline{\beta}_{i} & -\underline{\beta}_{s} \right) \end{bmatrix}^{T} \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$
$$\Rightarrow S = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \Rightarrow S_{\alpha} = \begin{bmatrix} a & -c \\ -b & d \end{bmatrix}$$
(2.97)

These are just two elements of the scattering matrix group identified in equation (1.145). As shown there, this can be interpreted in terms of the scatterer as a 180-degree rotation about the bisectrix \underline{b} , as shown in Figure 2.7. This we define as the scatterer being in its reciprocal position. Finally, we show the effect on



Fig. 2.7 Reciprocity interpreted as a rotation about the bisectrix

the associated scattering vectors in equation (2.98):

$$\underline{k} = \begin{bmatrix} a+d\\a-d\\b+c\\i(b-c) \end{bmatrix} \rightarrow \underline{k}_{\alpha} = \begin{bmatrix} a+d\\a-d\\-(b+c)\\i(b-c) \end{bmatrix}$$
(2.98)

Therefore, assuming a random collection of scatterers formed from a mixed population with equal numbers in their original and reciprocal positions, then $\langle [T] \rangle$ and $\langle [M] \rangle$ have a form predicted by adding (incoherently) the [T] and [M] matrices for S and S_{α} in equation (2.97) to yield the form shown in equation (2.99). This shows that in the bistatic case, (hv+vh) is uncorrelated with the copolar channels, but that (hv–vh) can maintain correlation. The impact of these correlations on the form of the Mueller matrix is shown on the right in equation (2.99):

$$\left\langle T_{FSA}^{reciprocity} \right\rangle = \begin{bmatrix} t_{11} & t_{12} & 0 & t_{14} \\ t_{12}^* & t_{22} & 0 & t_{24} \\ 0 & 0 & t_{33} & 0 \\ t_{14}^* & t_{24}^* & 0 & t_{44} \end{bmatrix} \Leftrightarrow \left\langle M_{FSA}^{reciprocity} \right\rangle$$

$$= \begin{bmatrix} m_{11} & m_{12} & m_{13} & m_{14} \\ m_{12} & m_{22} & m_{23} & m_{24} \\ -m_{13} & -m_{23} & m_{33} & m_{34} \\ m_{14} & m_{24} & -m_{34} & m_{44} \end{bmatrix}$$

$$(2.99)$$

Consequently, the polarised/depolarised decomposition for the reciprocal bistatic scattering case considerably simplifies. The 4×4 unitary matrix of eigenvectors can now be expressed as a 3×3 sub-unitary matrix, as shown in equation (2.100). In this case we return to the N = 3 scenario, and SU(2) governs the depolarisation subspace:

$$[U_4] = \begin{bmatrix} [U_3] & 0\\ 0 & 1 \end{bmatrix} \Rightarrow [U_{-1}] [U_4] = \begin{bmatrix} [U_2] & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1 \end{bmatrix}$$
(2.100)

However, there is now one extra depolarising eigenvalue to consider, and so the polarised/depolarised can be written in terms of a reduced set of ten parameters, as shown in equation (2.101):

$$T_4^{recip} = (4+1) + [2+3] \tag{2.101}$$

Following this logic there are two other symmetry combinations to consider from the scattering matrix group. In the first we combine a scatterer with its mirror image in the plane of scattering, defined from the incident and scattered wave vectors \underline{k} , as shown in Figure 1.14. The scattering vectors and averaged coherency and Mueller matrices for this case then have the form shown in

equation (2.102):

$$\underline{k} = \begin{bmatrix} a+d\\ a-d\\ b+c\\ i(b-c) \end{bmatrix} \rightarrow \underline{k}_{\beta} = \begin{bmatrix} a+d\\ a-d\\ -(b+c)\\ -i(b-c) \end{bmatrix}$$
$$\rightarrow \langle T \rangle = \begin{bmatrix} t_{11} & t_{12} & 0 & 0\\ t_{12}^* & t_{22} & 0 & 0\\ 0 & 0 & t_{33} & t_{34}\\ 0 & 0 & t_{34}^* & t_{44} \end{bmatrix} \Leftrightarrow \langle M \rangle = \begin{bmatrix} m_{11} & m_{12} & 0 & 0\\ m_{21} & m_{22} & 0 & 0\\ 0 & 0 & m_{33} & m_{34}\\ 0 & 0 & m_{43} & m_{44} \end{bmatrix}$$
(2.102)

This case is again governed by SU(4), but the 4×4 unitary matrix can be factored in terms of 2×2 unitary matrices, as shown in equation (2.103):

$$\begin{bmatrix} U_4 \end{bmatrix} = \begin{bmatrix} [U_2] & [0] \\ [0] & [U_2] \end{bmatrix} \Rightarrow \begin{bmatrix} U_{-1} \end{bmatrix} \begin{bmatrix} U_4 \end{bmatrix} = \begin{bmatrix} [I_2] & [0] \\ [0] & [U_2] \end{bmatrix}$$
(2.103)

where $[I_2]$ is the 2 × 2 identity matrix, and [0] the 2 × 2 null matrix. Again SU(2) governs the depolarisation subspace, and the polarised/depolarised decomposition has eight parameters organized as shown in equation (2.104):

$$T_4^{plane} = (2+1) + [2+3] \tag{2.104}$$

Finally we consider a combination of a scatterer and its mirror image in the bisectrix plane, defined orthogonal to the scattering plane and including the bisectrix vector \underline{b} . In this case the scattering vectors and averaged coherency and Mueller matrices have the form shown in equation (2.105):

$$\underline{k} = \begin{bmatrix} a+d\\ a-d\\ b+c\\ i(b-c) \end{bmatrix} \rightarrow \underline{k}_{\gamma} = \begin{bmatrix} a+d\\ a-d\\ b+c\\ -i(b-c) \end{bmatrix}$$
$$\rightarrow \langle T \rangle = \begin{bmatrix} t_{11} & t_{12} & t_{13} & 0\\ t_{12}^* & t_{22} & t_{23} & 0\\ t_{13}^* & t_{23}^* & t_{33} & 0\\ 0 & 0 & 0 & t_{44} \end{bmatrix} \Leftrightarrow \langle M \rangle = \begin{bmatrix} m_{11} & m_{12} & m_{13} & m_{14}\\ m_{12} & m_{22} & m_{23} & m_{24}\\ m_{13} & m_{23} & m_{33} & m_{34}\\ -m_{14} & -m_{24} & -m_{43} & m_{44} \end{bmatrix}$$
(2.105)

Again the SU(4) eigenvector dependence can be represented as a 3×3 unitary matrix as shown in equation (2.106), so again SU(2) controls the depolarisation subspace.

$$\begin{bmatrix} U_4 \end{bmatrix} = \begin{bmatrix} [U_3] & 0 \\ 0 & 1 \end{bmatrix} \Rightarrow \begin{bmatrix} U_{-1} \end{bmatrix} \begin{bmatrix} U_4 \end{bmatrix} = \begin{bmatrix} [U_2] & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$
(2.106)

The polarised/depolarised decomposition then has ten parameters arranged as shown in equation (2.107):

$$T_4^{bi \sec trix} = (4+1) + [2+3] \tag{2.107}$$

In the most general symmetry case, when all the above symmetries apply—as, for example, for a random cloud of spheroidal particles—we obtain the averaged bistatic coherency matrix by adding coherency matrices for *all* elements of the scattering matrix group, as shown in equation (2.108):

$$\underline{k} = \begin{bmatrix} a+d\\ a-d\\ b+c\\ i(b-c) \end{bmatrix} \rightarrow \underline{k}_{\alpha} = \begin{bmatrix} a+d\\ a-d\\ -(b+c)\\ i(b-c) \end{bmatrix} \rightarrow \underline{k}_{\beta} = \begin{bmatrix} a+d\\ a-d\\ -(b+c)\\ -i(b-c) \end{bmatrix} \rightarrow \underline{k}_{y} = \begin{bmatrix} a+d\\ a-d\\ b+c\\ -i(b-c) \end{bmatrix}$$
$$\Rightarrow \langle T_{wave}^{symmetric} \rangle = \begin{bmatrix} t_{11} & t_{12} & 0 & 0\\ t_{12}^{*} & t_{22} & 0 & 0\\ 0 & 0 & t_{33} & 0\\ 0 & 0 & 0 & t_{44} \end{bmatrix} \Leftrightarrow \langle M \rangle = \begin{bmatrix} m_{11} & m_{12} & 0 & 0\\ m_{12} & m_{22} & 0 & 0\\ 0 & 0 & m_{33} & m_{34}\\ 0 & 0 & -m_{34} & m_{44} \end{bmatrix}$$
(2.108)

This is the bistatic equivalent of azimuthal symmetry in backscatter problems. We see that in this case the SU(4) matrix governing the eigenvectors reduces to an SU(2) dependency as shown in equation (2.109), and so again there are no eigenvector parameters in the depolarisation subspace.

$$\begin{bmatrix} U_4 \end{bmatrix} = \begin{bmatrix} [U_2] & [0] \\ [0] & [I_2] \end{bmatrix} \Rightarrow \begin{bmatrix} U_{-1} \end{bmatrix} \begin{bmatrix} U_4 \end{bmatrix} = \begin{bmatrix} I_3 \end{bmatrix}$$
(2.109)

In this case the polariser/depolariser decomposition takes on the reduced form shown in equation (2.110):

$$[T] \xrightarrow{bistatic+symmetry} (2+1) + [0+3]$$
(2.110)

2.4.3.2 The scattering sphere

This is a very useful result, as this type of symmetry is quite common in environmental remote sensing applications (see Chapter 3). Here we see that in the most symmetric case, the polarising/depolarising decomposition can be characterized by six parameters in a 2:1:0:3 cascade. Hence the principal eigenvector can be entirely associated with polarised behaviour of the system, with a scattering strength given by the largest eigenvalue of [*T*]. The principal eigenvector has only two complex elements, and hence corresponds to diagonal amplitude matrices. Now using the SU(2)–O3 mapping (see Appendix 2) we can project all such diagonal matrices onto the surface of a sphere in a real three-dimensional space. This sphere shares many properties with the Poincaré sphere, but each point now represents a scattering amplitude matrix rather than a wave state. To distinguish the two we term this a scattering sphere.

A general polarised eigenvector can then be parameterized in terms of two angles on this sphere, as shown in equation (2.111) and geometrically in Figure 2.8.

 $[S] = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ Scattering sphere

Fig. 2.8 Scattering sphere showing the geometrical interpretation of α and δ from equation 2.111

$$(2+1) \rightarrow \sqrt{\lambda_{\max}} \begin{bmatrix} \cos \alpha \\ \sin \alpha e^{i\delta} \end{bmatrix}$$
 (2.111)

Note that $\alpha = \delta = 0$ corresponds to the identify amplitude matrix, and so any departure of these parameters from this point represents a departure from trivial scattering behaviour. There also remain three sub-eigenvalues that fully characterize the types of depolarisation that can occur around this polarised system.

While it is possible to classify directly in the space of δ_1 , δ_2 , δ_3 , it is often again more convenient to use secondary parameters derived from the eigenvalues and more closely related to the degree of polarisation widely used in Stokes algebra. These parameters are defined in equation (2.112) as the scattering entropy *H* and scattering anisotropies A_{ij} :

$$H = -\sum_{i=1}^{4} P_i \log_4 P_i \quad 0 \le H \le 1, \quad P_i = \frac{\lambda_i}{\sum \lambda}$$

$$A_{ij} = \frac{\lambda_i - \lambda_j}{\lambda_i + \lambda_j} \quad 0 \le A_{ij} \le 1, \quad i > j$$
(2.112)

All vary between 0 and 1, with entropy being a general measure of total depolarisation. It is zero for non-depolarising systems, and 1 for the ideal depolariser. The anisotropies then provide information about the variation of depolarisation with changes in polarisation: if they are zero then there is no variation, while as they approach 1 there exists subspaces in the polarisation domain where depolarisation can be small. This then leads us naturally back to the idea of the entropy/alpha decomposition.

2.4.3.3 The bistatic entropy/alpha decomposition

For backscatter problems we found that the simplified two-parameter characterization of the entropy/alpha plane provided a convenient way to map different types of depolarisation. One attractive feature of this approach is that it scales to different dimensions, including the N = 4 bistatic case. In this case the change of dimensionality is accommodated by an extra eigenvalue and eigenvector, as shown in equation (2.113) (Cloude, 2005b, 2006a):

$$\langle T \rangle = U_4 \begin{bmatrix} \lambda_1 & 0 & 0 & 0\\ 0 & \lambda_2 & 0 & 0\\ 0 & 0 & \lambda_3 & 0\\ 0 & 0 & 0 & \lambda_4 \end{bmatrix} U_4^{*T} \quad U = \begin{bmatrix} \underline{e}_4^1 & \underline{e}_4^2 & \underline{e}_4^3 & \underline{e}_4^4 \end{bmatrix}$$

$$\underline{e}_4 = \begin{bmatrix} \cos \alpha \ e^{i\phi_1} \\ \sin \alpha \ \sin \psi \ \cos \gamma \ e^{i\phi_2} \\ \sin \alpha \ \sin \psi \ \sin \gamma \ e^{i\phi_4} \end{bmatrix}$$

$$(2.113)$$

The entropy/alpha parameters can now be defined as shown in equation (2.114):

$$H = -\sum_{i=1}^{4} P_i \log_4 P_i \quad 0 \le H \le 1, \quad \overline{\alpha} = \sum_{i=1}^{4} P_i \alpha_i \quad 0 \le \overline{\alpha} \le 90^\circ \quad (2.114)$$

Again there are bounding curves in this plane. The upper and lower values for $\overline{\alpha}$ as a function of increasing entropy can be calculated as shown in equation (2.115). Figure 2.9 shows how the valid region in the H/ α plane for N = 4 bistatic scattering differs from that used in the N = 3 backscatter and N = 2 dual polarisation cases.

At low entropies there is little difference between the N = 2 and N = 3 cases, but at higher entropy values the alpha range for N = 4 is shifted to higher values. This is a consequence of the addition of a new eigenvector, itself having an alpha value of $\pi/2$. This acts to lift the mean alpha value for any given entropy. In the limit of H = 1, the alpha range is reduced to a single point at $3\pi/8$. Again the utility of this diagram is not so much in its high fidelity parameterization of depolarisation in N = 4 cases, but in the convenience of representing the most general bistatic scattering problem by one polarising and one depolarising parameter, in such a way as to maintain invariance to changes



Fig. 2.9 Summary of entropy/alpha diagrams for N = 2, 3, and 4 scattering

of polarisation base.

$$T_{I} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & m & 0 & 0 \\ 0 & 0 & m & 0 \\ 0 & 0 & 0 & m \end{bmatrix} \Rightarrow \begin{cases} \overline{\alpha} = \frac{3m\pi}{2(1+3m)} \\ H = -\frac{1}{1+3m} \log_{4} \frac{m^{3m}}{(1+3m)^{3m-1}} \\ \end{bmatrix}$$

$$\begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 2m & 0 \\ 0 & 0 & 0 & 2m \end{bmatrix} \quad 0 \le m \le 0.5 \quad \Rightarrow \begin{cases} \overline{\alpha} = \frac{\pi}{2} \\ H = -\frac{1}{(1+4m)} \log_{4} \frac{2m^{4m}}{(1+4m)^{4m+1}} \\ \end{bmatrix}$$

$$\begin{bmatrix} 2m-1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad 0.5 \le m \le 1 \quad \Rightarrow \begin{cases} \overline{\alpha} = \frac{3\pi}{4m+4} \\ H = -\frac{1}{(2+2m)} \log_{4} \frac{(2m-1)^{2m-1}}{(2+2m)^{2+2m}} \\ \end{bmatrix}$$

$$(2.115)$$

2.5 Characterization of depolarising systems

Finally we amalgamate the results of the previous sections to re-examine the idea of depolarisation and its various parameterizations. For example, we saw that the vector wave reciprocity theorem in backscatter causes a symmetry in [S] which limits the form of the Mueller matrix (for arbitrary random scattering problems) to that shown in equation (2.116), where we note that there is an important additional constraint equation on the diagonal elements, leaving [M] with only nine rather than sixteen degrees of freedom. Reciprocity symmetry therefore acts to limit the types of depolarisation we can observe in backscatter.

$$[S]_{FSA} = \begin{bmatrix} a & b \\ -b & d \end{bmatrix} \Rightarrow [M]_{FSA} = \begin{bmatrix} m_{11} & m_{12} & m_{13} & m_{14} \\ m_{12} & m_{22} & m_{23} & m_{24} \\ -m_{13} & -m_{23} & m_{33} & m_{34} \\ m_{14} & m_{24} & -m_{34} & m_{44} \end{bmatrix}$$
(2.116)

Reciprocity $\Rightarrow m_{11} - m_{22} + m_{33} - m_{44} = 0$

The most extreme example is the isotropic depolariser, with a Mueller matrix $[M_I]$ of the form shown on the left-hand side of equation (2.117):

This matrix converts all Stokes vectors into a randomly polarised wave, but there is no corresponding single [S] matrix. This form leads to a standard

generalization of the depolariser, as shown in two stages from left to right in equation (2.117). The middle $[M_{II}]$ is a partial depolariser, while the right-hand form $[M_{III}]$ generalizes to an anisotropic partial depolariser with arbitrary direction in Stokes space (the matrix O_3 is a 3×3 real rotation matrix of the Poincaré sphere). As an example of a partial depolariser, consider the special case of forward scattering by random particles that takes the form shown in equation (2.118):

$$[M] = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \delta & 0 & 0 \\ 0 & 0 & \delta & 0 \\ 0 & 0 & 0 & \varepsilon \end{bmatrix} \Rightarrow \begin{cases} |\varepsilon| \le 1 \\ |\delta| \le 1 \\ 2\delta - \varepsilon \le 1 \end{cases}$$
(2.118)

Again we note important physical restrictions on the diagonal elements from the forward scattering symmetry. Given the need to always add additional constraint equations, we can ask if all triplets $\delta_1, \delta_2, \delta_3$ are physically consistent, and whether the forms in equation (2.117) exhaust all possibilities.

In this regard we have shown how we can use the coherency matrix concept to classify depolarisers. For example, by mapping the general depolariser Δ of equation (2.117) into [*T*] we see that the real diagonal elements δ_1, δ_2 and δ_3 are constrained by the four inequalities shown in equation (2.119). We can interpret this geometrically by first considering a real three-dimensional space formed by triplets $\delta_1, \delta_2, \delta_3$. If we normalize the Mueller matrix to have $m_{11} =$ 1, the unit cube in this space then represents the set of candidate depolarisers, as shown on the left-hand side of Figure 2.10. Equation (2.119) then represents the equation of four planes in this space, and all physical depolarisers must be inside the volume bounded by these planes. The shape of this region is shown on the right in Figure 2.10. To illustrate the sometimes subtle implications of these constraints, we show, in equation (2.120), three candidate depolarisers, all of which seem to be reasonable suggestions:



Fig. 2.10 Section of the depolarising unit cube (left) and restricted physical depolariser region (right)

$$[M_1] = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0.5 & 0 & 0 \\ 0 & 0 & 0.4 & 0 \\ 0 & 0 & 0 & 0.3 \end{bmatrix} [M_2] = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0.5 & 0 & 0 \\ 0 & 0 & 0 & -0.3 \end{bmatrix}$$
$$[M_3] = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0.5 & 0 & 0 \\ 0 & 0 & -0.4 & 0 \\ 0 & 0 & 0 & -0.3 \end{bmatrix}$$
(2.120)

However, using the constraints of equation (2.119) we see that $[M_2]$ is not a physical depolariser. Its corresponding coherency matrix has a negative eigenvalue, and hence no physical system can generate such a depolarisation process. Note, however, that $[M_1]$ and $[M_3]$ are both located inside the feasible volume and therefore correspond to realizable depolarisers.

The constraints on forward scattering and backscattering matrices in equations (2.116) and (2.118) can now also be seen as special cases of this geometry. In forward scattering, rotation symmetry constrains $\delta_1 = \delta_2 = \delta$, and the following special case is thus obtained:

We see that these inequalities give rise to the constraints shown in equation (2.118). Equally in backscatter, the reciprocity theorem reduces [S] by relating the crosspolarised elements, and this has the effect of reducing the rank of [T] from 4 to 3. Hence one eigenvalue (λ_3) of [T] must always be zero, even in the presence of depolarisation. This then corresponds directly to the constraint equation shown in equation (2.116). We see that the geometry of Figure 2.10 is important in classifying depolarisers, and underlines the more fundamental role played by [T] rather than [M] in understanding processes of wave scattering and depolarisation.

We can further extend this analysis to include the case of a 'general' depolariser proposed on the far right in equation (2.117), and obtained by a rotation of the Poincaré sphere (Lu, 1996). The transformation of [T] under such a rotation can be derived using the mapping from rotations in a real threedimensional space to unitary transformations in a two-dimensional complex space: the SU(2)– O_3 homomorphism, as shown in equation (2.122).

$$\underline{g'} = \begin{bmatrix} 1 & 0 \\ 0 & [O_3] \end{bmatrix} \underline{g} \xrightarrow{SU(2) - O_3^+} [S]' = [U_2] [S] [U_2]^{*T}$$

$$\Rightarrow \langle [T] \rangle' = [U_{4B}] \langle [T] \rangle [U_{4B}]^{*T} [U_{4B}] = [U_2] \otimes [U_2]^{*}$$
(2.122)

We see now, however, that this represents only a subset of possible depolarisers, and a more general classification should be based on the full 4×4 unitary matrix transformations of $[T]' = [U_4][T][U_4]^{*T}$. This underlines the importance of general unitary transformations in the analysis of polarisation phenomena.

2.6 Relating the Stokes/Mueller and coherency matrix formulations

In this chapter we have introduced an important new formulation of polarised scattering—the coherency matrix [T]—and it is now of interest to formally connect the coherency and Mueller matrix formulations via their treatment of scattered intensity. We start by again considering a Pauli matrix expansion of the amplitude matrix as shown in equation (2.123):

$$\underline{E}_{s} = [S] \underline{E}_{i} = \frac{1}{\sqrt{2}} \begin{bmatrix} k_{0} + k_{1} & k_{2} + ik_{3} \\ k_{2} - ik_{3} & k_{0} - k_{1} \end{bmatrix} \underline{E}_{i} \quad k_{0}, k_{1}, k_{2}, k_{3} \in C$$
(2.123)

We then consider a hypothetical (but always physically realizable) measurement system, represented now by a complex four-element vector \underline{w} , formed as complex weights of the Pauli coefficients, as shown in equation (2.124):

$$p = w_1 k_0 + w_2 k_1 + w_3 k_2 + w_4 k_3 = \underline{w}^{*T} \underline{k} \in C$$

$$|p|^2 \ge 0 \quad for \quad \underline{w} \in C_4 \quad \underline{w}^{*T} \underline{w} = 1$$
(2.124)

In particular we note that for arbitrary choice of \underline{w} , the squared amplitude of the projection *p* is always non-negative, as shown. This is important when we consider generalization to the case of scattering from random media, when equation (2.224) takes on the form shown in equation (2.125), where $\langle .. \rangle$ is now an ensemble average.

$$m = \langle pp^{*T} \rangle = \langle \underline{w}^{*T} \underline{k} \underline{k}^{*T} \underline{w} \rangle = \underline{w}^{*T} \langle \underline{k} \underline{k}^{*T} \rangle \underline{w} = \underline{w}^{*T} [T] \underline{w} \ge 0 \qquad (2.125)$$

Here [*T*] is the 4 × 4 scattering coherency matrix. We see that the non-negative constraint from equation (2.124) requires [*T*] to be positive semidefinite; that is, to have a non-negative real eigenvalue spectrum. Now we have seen that [*T*] is related in a 1–1 mapping to the 4 × 4 real Mueller matrix [*M*]. The explicit mapping from [*M*] into [*T*] is shown in equations (2.48) and (2.49). More formally, we can relate the sixteen-element vectors \underline{t} and \underline{m} , formed by expanding [*T*] and [*M*] row-wise: $\underline{m} = [m_{11}, m_{12}, m_{13} \dots m_{44}]^T$, by a 16 × 16 complex matrix [*Q*], as shown in equation (2.126). This gives the same result as in equation (2.48), but with a more formal notation we can make use of further analysis:

$$\underline{t} = [Q] \underline{m} \Rightarrow [Q] = \frac{1}{2} \begin{bmatrix} Q_1 & Q_2 & Q_3 & Q_4 \\ Q_2 & Q_1 & iQ_4 & -iQ_3 \\ Q_3 & -iQ_4 & Q_1 & iQ_2 \\ Q_4 & iQ_3 & -iQ_2 & Q_1 \end{bmatrix}$$

$$Q_1 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} Q_2 = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & i \\ 0 & 0 & -i & 0 \end{bmatrix}$$

$$Q_3 = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -i \\ 1 & 0 & 0 & 0 \\ 0 & i & 0 & 0 \end{bmatrix} Q_4 = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & i & 0 \\ 0 & -i & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}$$
(2.126)

For example, we can now relate the scattered intensity in the Stokes/Mueller formulation to that in the coherency matrix form by equating expressions for scattered intensity and using [Q], as shown in equation (2.127):

$$\underline{w}^{*T}[T]\underline{w} = \underline{g}_{r}^{T}[M] \underline{g}_{t} \Rightarrow (\underline{w}^{*T} \otimes \underline{w}^{T}) \underline{t} = (\underline{g}_{r}^{T} \otimes \underline{g}_{t}^{T}) \underline{g}$$

$$\Rightarrow (\underline{w}^{*T} \otimes \underline{w}^{T}) [Q] = (\underline{g}_{r}^{T} \otimes \underline{g}_{t}^{T})$$

$$\Rightarrow (\underline{w}^{*T} \otimes \underline{w}^{T}) = (\underline{g}_{r}^{T} \otimes \underline{g}_{t}^{T}) [Q]^{-1}$$
(2.127)

Hence any choice of transmit and receive Stokes vectors g_t and g_r can be transformed into an equivalent w formulation and *vice versa*. Given a complex w we can convert it into a weight vector for the elements of [M], a subset of which can then be further decomposed into a direct product of Stokes vectors. Note, however, that from a physical point of view the most important relationship is equation (2.124). This relationship is not so evident in the Stokes/Mueller formulation, and so we can in some sense consider the [T] formulation to be more fundamental.

In the previous chapters we considered in detail an algebraic formulation of vector wave propagation and scattering, and have used this to establish a detailed parameterization of depolarisation phenomena. It is now time to consider examples of the application of this theory to physical models for surface and volume scattering.

Depolarisation in surface and volume scattering

3

In Chapter 1 we introduced the scattering amplitude matrix [S] and its transformation properties, and in Chapter 2 we saw how to extend these concepts to include the presence of depolarisation. Here we derive examples of the explicit form of the amplitude matrix and its coherency structure for two important canonical problems: surface scattering and volume scattering. Details of the underlying derivations are widely available in the existing literature, with several excellent books covering the topics required (Born and Wolf, 1989; Ishimaru, 1991; Jackson, 1999; Kong, 1985; Lakhtakia, 1989; Tsang, 1985; Ulaby, 1986; van de Hulst, 1981; Mishchenko, 2000). Here our purpose is mainly to concentrate on the polarimetric properties of these solutions, and so we proceed by first summarising the key formulae and then move on to highlight their important polarimetric structure in terms of polarised and depolarised component decompositions. We also concentrate on those models that a) provide a *fully* polarimetric solution, including phase and amplitude effects, and b) lead to simple analytical results without the need for numerical methods such as finite difference and finite element. This will enable us to establish a few suitable benchmarks for later comparison with numerical experiments or more advanced modelling techniques. Inevitably this means that we concentrate on low-frequency solutions, where analysis is tractable. Hence we start by considering scattering by electrically smooth surfaces and by volumes composed of electrically small particles. However, these not only have direct relevance to low-frequency radar applications (particularly at L and P bands), but also provide a framework by which to judge more complicated high-frequency solutions.

In order to calculate a general [S] matrix we must solve Maxwell's differential equations together with appropriate boundary conditions applied on the surface of the object. From the former we postulate polarised plane wave solutions and, by enforcing the latter, obtain a set of equations for the unknown co- and crosspolarised scattering coefficients in a chosen coordinate system. These can then be solved for two orthogonal incident polarisation states to obtain the four complex elements of the scattering amplitude matrix [S]. Key to this process is an understanding of the behaviour of vector EM fields across a boundary between two materials.

There are four sets of boundary conditions arising from the set of four Maxwell equations (see equation (1.1)). At an interface between two media, the two curl equations require continuity of the tangential components of \underline{B} and \underline{H} , while the divergence equations require continuity of the normal components of \underline{D} and \underline{B} . For a general boundary we then define the local normal vector \underline{n} , as shown in Figure 3.1. Note that of the four sets of boundary conditions only two

$$\underbrace{\underline{E}_{1} \underline{H}_{1} \underline{D}_{1} \underline{B}_{1}}_{\underline{E}_{2} \underline{H}_{2} \underline{D}_{2} \underline{B}_{2}} \underbrace{\underline{n} \times \underline{E}_{1} = \underline{n} \times \underline{E}_{2}}_{\underline{n} \times \underline{H}_{1} = \underline{n} \times H_{2}} \underbrace{\underline{n} . \underline{D}_{1} = \underline{n} . \underline{D}_{2}}_{\underline{n} \times \underline{H}_{1} = \underline{n} \times H_{2}} \underbrace{\underline{n} . \underline{D}_{1} = \underline{n} . \underline{D}_{2}}_{\underline{n} \times \underline{H}_{1} = \underline{n} \times H_{2}}$$

Fig. 3.1 Vector boundary conditions in electromagnetic scattering are independent, the others being related by Maxwell's equations. In practice, therefore, any two of the four can be chosen, depending on analytical convenience. We now turn to consider the application of these ideas to the first of our two themes: surface scattering.

3.1 Introduction to surface scattering

As a simple example of the above procedure, consider the classical problem of reflection and transmission at a plane dielectric interface (Born and Wolf, 1989; Hecht and Zajac, 1997). The upper medium has a relative dielectric constant ε_1 and the lower ε_2 , as shown in Figure 3.2. On the left-hand side is shown the case for an incident electric field polarised parallel to the plane of incidence. This corresponds to vertical polarisation, V, in radar applications, otherwise termed 'p' polarisation or TM for transverse magnetic wave. The magnetic field vector H is shown at right angles to both the electric field and the wave propagation vector β_{i} , as follows from Maxwell's equations for wave propagation. From the laws of reflection and refraction, the reflected and transmitted fields are zero everywhere except in specific directions. The reflected wave is such that the surface normal is the bisectrix vector (angle of incidence equals the angle of reflection), and the refracted wave obeys Snell's law relating the transmission angle to the incident angle and refractive index n, given by the Maxwell relation as the square root of dielectric constant, as shown in equation (3.1) (Born and Wolf, 1989; Hecht and Zajac, 1997):

$$n_1 \sin \theta_i = n_2 \sin \theta_t \, n = \sqrt{\varepsilon_r} \tag{3.1}$$

From equation (1.145) we can then generate the S_{β} daughter problem (without yet calculating the detailed form of the matrix), which from the symmetry of the plane must be equal to the original mother matrix S. The only solution to this constraint is that b = c = 0, or the matrix has zero crosspolarisation. Hence a smooth infinite plane interface causes zero crosspolarisation. This leaves only the two diagonal copolar elements of the amplitude matrix. Note that this is true for any scattering angle θ (not just specular) and for arbitrary surface roughness, as long as the roughness profile is one-dimensional (like a corrugated rough surface).

For each polarisation combination, such as VV, we have two unknowns: the reflected and transmitted electric field components. However, by matching



Fig. 3.2 Surface reflection and transmission for incident V (parallel, 'p' or TM) and H (perpendicular, 's' or TE) polarisations

V or parallel (p) incident polarisation H o

H or perpendicular (s) incident polarisation

the tangential components of \underline{E} and \underline{H} across the boundary, two independent equations can be obtained. This is sufficient to provide a full solution, as we now demonstrate. First, however, we consider the perpendicular polarisation case.

On the right-hand side of Figure 3.2 we show the corresponding HH problem (often called 's' polarisation from the German *senkrecht*, meaning perpendicular, or alternatively TE for Transverse Electric waves) when the electric field is now polarised perpendicular to the plane of incidence. The same procedure can again be used to find the reflection and transmission coefficients, as follows.

We start by defining the values of the field components at the interface (z = 0) by subscripts 'r' for reflected, 'i' for incident, and 't' for transmitted. We then choose to enforce continuity of tangential *E* and *H* fields (see Figure 3.1) to obtain two equations for the unknown coefficients, as shown in equation (3.2):

$$H_r^s \cos \theta_r - H_i^s \cos \theta_i = -H_t^s \cos \theta_t \Rightarrow (H_r^s - H_i^s) \cos \theta_i = -H_t^s \cos \theta_t$$
$$E = \frac{c\mu_o}{n} H \Rightarrow \begin{cases} E_i^s + E_r^s = E_t^s \\ n_1(E_i^s - E_r^s) \cos \theta_i = n_2 E_t^s \cos \theta_t \end{cases}$$
(3.2)

The first equation follows from the continuity of tangential electric field for 's' polarisation. The second requires some manipulations, projecting the H field onto the x axis and using the relationship between E and H components of a plane wave from Maxwell's equations. Note that c is the velocity of light in free space.

Similarly, for the V or 'p' polarised problem, when the *E* field is polarised in the plane of incidence, we can obtain two equations by enforcing continuity of tangential *E* and *H*, as shown in equation (3.3):

$$\begin{cases} n_1(E_i^p + E_r^p) = n_2 E_t^p \\ (E_i^p - E_r^p) \cos \theta_i = E_t^p \cos \theta_t \end{cases}$$
(3.3)

Both solutions for reflection and transmission coefficients can then be combined into the reflection and transmission matrices, as shown in equation (3.4). Note that for simplicity of notation we assume non-magnetic materials where the relative permeability $\mu_r = 1$.

3.1.1 The Fresnel equations

$$[R] = \begin{bmatrix} R_{HH} & 0\\ 0 & R_{VV} \end{bmatrix} \Rightarrow \begin{cases} R_{HH} = \frac{n_1 \cos \theta_i - n_2 \cos \theta_i}{n_1 \cos \theta_i + n_2 \cos \theta_i}\\ R_{VV} = \frac{n_2 \cos \theta_i - n_1 \cos \theta_i}{n_2 \cos \theta_i + n_1 \cos \theta_i} \end{cases}$$
(3.4)
$$[T] = \begin{bmatrix} T_{HH} & 0\\ 0 & T_{VV} \end{bmatrix} \Rightarrow \begin{cases} T_{HH} = \frac{2n_1 \cos \theta_i}{n_1 \cos \theta_i + n_2 \cos \theta_i}\\ T_{VV} = \frac{2n_1 \cos \theta_i}{n_2 \cos \theta_i + n_1 \cos \theta_i} \end{cases}$$

These four expressions are called the Fresnel equations, after the French scientist Augustin-Jean Fresnel (1788–1827). Despite their relative simplicity and longevity, they remain of fundamental importance in understanding the interaction of polarised waves with media. The reflection coefficients can also be expressed in terms of 'exterior' parameters θ_i and ε_2 by using Snell's law (equation (3.1)) to remove θ_2 , as shown in equation (3.5):

$$R_{HH} = \frac{\cos\theta_i - \sqrt{\frac{\varepsilon_2}{\varepsilon_1} - \sin^2\theta_i}}{\cos\theta_i + \sqrt{\frac{\varepsilon_2}{\varepsilon_1} - \sin^2\theta_i}} \quad R_{VV} = \frac{\frac{\varepsilon_2}{\varepsilon_1}\cos\theta_i - \sqrt{\frac{\varepsilon_2}{\varepsilon_1} - \sin^2\theta_i}}{\frac{\varepsilon_2}{\varepsilon_1}\cos\theta_i + \sqrt{\frac{\varepsilon_2}{\varepsilon_1} - \sin^2\theta_i}} \quad (3.5)$$

This latter form is more commonly used in radar scattering problems. Furthermore, the incident wave is often in free space, and so we can set $\varepsilon_1 = 1$ to further simplify the equations. Note that these equations apply even for complex permittivity ε_2 , such as occurs for microwave reflection from wet soils and sea ice, for example. In this case the wave is changed in both amplitude and phase on reflection, but equation (3.5) is still valid. We now turn to consider the properties of such lossy material in more detail.

3.1.1.1 Fresnel equations for lossy dielectric media

In the presence of losses, the dielectric constant of the lower medium must be written (for an exp(i ω t) time variation) as a complex number $\varepsilon_r = \varepsilon' - i\varepsilon''$, where the ratio tan $\delta = \frac{\varepsilon''}{\varepsilon'}$ is called the loss tangent of the material. For example, in a conducting medium with conductivity σ it follows from Maxwell's equations that $J = \sigma E$ and the dielectric constant becomes complex with wavenumber β , as shown in equation (3.6):

$$\varepsilon_r = \varepsilon' - i\varepsilon'' \approx \varepsilon' - i\frac{\sigma}{\omega\varepsilon_o} \Rightarrow \beta = \omega_0\sqrt{\mu\varepsilon} = \beta_0 n = \beta_0 n' - i\beta_0 n'' \quad (3.6)$$

Note, however, that since the material remains homogeneous, the wave polarisation state does not change with propagation into the material. The only 'filtering' of relative polarisation amplitude and phase is by the surface transmission given by the Fresnel equations (3.4). However, it is of interest to study a little further the properties of waves propagating in such lossy material. This will lead us to some important concepts concerning penetration depth and its relationship to volume scattering.

To study the propagation of signals into lossy material we first write the wave vector in the material in terms of its 'x' (parallel to the interface) and 'z' (perpendicular to the interface) components, as shown in equation (3.7):

$$E_2 = T(\theta) \exp(-i(\beta_{2x}x + \beta_{2z}z))$$
(3.7)

where $T(\theta)$ is the appropriate Fresnel transmission coefficient. We are interested to see how such a wave propagates in the medium below the interface. The components of the wave vector can be explicitly derived as shown in equation (3.8), where we have assumed that the upper medium is free space and has propagation constant $\beta_1 = \beta_0 = \frac{2\pi}{\lambda}$.

$$\beta_{2x} = \beta_{1x} = \beta_0 \sin \theta_i$$

$$\beta_{2z} = \beta_2 \cos \theta_t = \sqrt{\beta_2^2 - \beta_1^2 \sin^2 \theta_i}$$

$$= \beta_0 \sqrt{\varepsilon' - i\varepsilon'' - \sin^2 \theta_i} = \beta_0 \sqrt{a + ib} = \beta_z - i\kappa_z$$
(3.8)

Evaluating the square root leads to the following result:

$$E_t = T \exp\left(-i\beta_0 \sin \theta_i x - i\beta_z z\right) \exp(-\kappa_z z)$$
(3.9)

where the components κ_z and β_z are real and defined, as shown in equation (3.10):

$$\beta_{z} = \frac{\beta_{0}}{\sqrt{2}} \sqrt{\varepsilon' - \sin^{2} \theta_{i}} \sqrt{\left(1 + \frac{\varepsilon''^{2}}{(\varepsilon' - \sin^{2} \theta_{i})^{2}}\right)^{0.5} + 1}$$

$$\kappa_{z} = \frac{\beta_{0}}{\sqrt{2}} \sqrt{\varepsilon' - \sin^{2} \theta_{i}} \sqrt{\left(1 + \frac{\varepsilon''^{2}}{(\varepsilon' - \sin^{2} \theta_{i})^{2}}\right)^{0.5} - 1}$$
(3.10)

This shows that the field inside the material decays exponentially from the surface (and to first order the decay rate is proportional to frequency, so increasing at higher frequencies) with coefficient κ_z , and that the equiamplitude and equiphase contours are no longer the same. This is called an *inhomogeneous* plane wave, and is characteristic of propagation into lossy material (those with a loss tangent greater than zero). This idea of projecting the wave vector $\underline{\beta}_o$ onto surface components is also an important one, and will be used extensively in our analysis of interferometry in Chapter 5.

Note that wave attenuation is more often expressed in terms of a *power* attenuation coefficient σ_e with units of decibels/meter (dB/m). Equation (3.11) can then be used for the conversion from κ_z .

$$\sigma_e^{dB} = \frac{20}{\ln(10)} \kappa_z = 8.686 \kappa_z \, dB/m$$

$$\kappa_z = \frac{\ln(10)}{20} \sigma_e^{dB} = 0.115 \sigma_e^{dB} \, m^{-1}$$
(3.11)

A key concept in such material is the penetration or skin depth, defined as the depth at which the signal is attenuated to exp(-1) (-8.686 dB). This can be considered a typical penetration depth into the material. In general, therefore, the penetration depth is given in terms of the complex refractive index, as shown in equation (3.12):

$$\delta = \frac{1}{\kappa_z} = \frac{\sqrt{2}}{\beta_0 \sqrt{\varepsilon' - \sin^2 \theta_i} \sqrt{\left(1 + \frac{\varepsilon''^2}{(\varepsilon' - \sin^2 \theta_i)^2}\right)^{0.5} - 1}}$$
(3.12)

This is the most general expression, but in the special case of very good conductors with conductivity σ this expression takes the more familiar form of the skin depth shown in equation (3.13):

$$\delta \approx \sqrt{\frac{2}{\omega\mu_0\sigma}} \tag{3.13}$$

As an important example of a lossy dielectric in microwave remote sensing, and one that illustrates the underlying complexity of modelling dielectric constant, we now summarize the main dielectric properties of soil.

3.1.1.2 Example: dielectric properties of soil

The interpretation of wave reflections from land surfaces involves an understanding of the dielectric constant of soil. This is such a basic requirement that it is perhaps surprising to realize the complexity of this problem. At microwave frequencies this complexity arises predominantly from the different forms that water can take in soil, and its impact on the resulting loss-tangent.

In general there are two main contributions to the loss-tangent of a material. The conductivity σ tends to be more important at low frequencies, but at higher frequencies dipolar losses due to molecular absorption tend to dominate, especially in the microwave spectrum. The predominant effect here, therefore, is water absorption, which has a broad resonant spectrum in the microwave region. Hence the presence of increasing water content tends to increase the loss tangent and thus reduce the penetration depth into the soil. However, complexity starts to arise when we consider that water in soil can be found in two main forms, bound or free, and these have very different dielectric properties. Since the proportion of bound and free water depends on soil structure, it follows that dielectric constant then depends on soil texture, and also shows a threshold effect; that is, above a certain transition water content the dielectric constant suddenly increases. In general, therefore, for lossy soils the loss tangent tends to be large at low frequencies and reduces to a minimum around 100 MHz before increasing again to reach a plateau around 3 GHz.

In more quantitative terms, soil can be considered a complex four-phase composite of air, bulk soil, and bound and free water, with a complex dielectric constant given by composite models of the general form shown in equation (3.14) (Dobson, 1985), which is valid in the frequency range 1.4–18 GHz.

$$\varepsilon_{soil}^{x} = 1 + \frac{\rho_b}{\rho_{ss}} \left(\varepsilon_{ss}^{x} - 1 \right) + m_v^y \varepsilon_{fw}^x - m_v \tag{3.14}$$

where the following terms can be defined:

- $\begin{aligned} \rho_b &= \text{soil bulk density (soil/air mixture)} \\ \rho_{ss} &= \text{density of solid soil (around 2.66 g/cm^3)} \\ \varepsilon_{ss} &= 4.7 i0 = \text{dielectric constant of dry soil} \\ x &= 0.65 \text{ (an average over 500 samples of five soil types)} \\ y &= y_0 y_1 S + y_2 C, S = \text{Sand fraction, } C = \text{Clay fraction, where the three real coefficients } y_0, y_1 \text{ and } y_2 \text{ are chosen to accommodate changes in soil texture (see Dobson, 1985)} \end{aligned}$
- m_v = volumetric moisture content

The imaginary component arises from the final term, defined from a Debye model of water absorption as follows:

$$\varepsilon_{fiw} = 4.9 + \frac{74.1}{1 + i\frac{f}{f_0}} - \frac{i\sigma_{\delta}}{2\pi\varepsilon_0 f}$$
 (3.15)

where σ_{δ} is the soil conductivity, and f_0 the relaxation frequency (which varies from 9 GHz at 0°C to 19 GHz at 25°C). Note that one of the main contributions to soil conductivity is its salinity 's_a' in ppthou with a typical derived relationship



Fig. 3.3 Vertical structure function for wave penetration into lossy material

of the form shown in equation (3.16) (Ulaby, 1986).

$$\sigma_{\delta} = 0.18252s_a - 1.4619e - 3s_a^2 + 2.093e - 5s_a^3 - 1.282e - 7s_a^4 \quad (3.16)$$

Typically, in remote sensing one is interested in estimating soil salinity or moisture indirectly from estimates of the dielectric constant [Oh, 1992; Dubois, 1995; Hajnsek, 2003; Allain, 2003). For example, a commonly used inversion formula for soil moisture is shown in equation (3.17) (Topp, 1980):

$$m_{\nu} = 10^{-2} (-5.3 + 2.92\varepsilon_r - 0.055\varepsilon_r^2 + 0.0004\varepsilon_r^3)$$
(3.17)

Hence we see that key to many remote sensing applications is estimation of dielectric constant, and later we shall see that polarisation diversity often provides a useful way to isolate dielectric constant in scattering problems.

In any case, we see that such lossy materials are characterized by exponential decay normal to the surface, as shown schematically in Figure 3.3. In this sense all homogeneous lossy media can be characterized by an exponential amplitude function with depth, with a characteristic penetration depth given by δ . We shall see later (in Chapter 8) that inhomogeneous materials can be characterized by different structure functions, and how in remote sensing we can devise methods for reconstruction of this function from scattered field data.

Finally, we note an interesting limit when κ_z tends to infinity and the penetration depth tends to zero. This occurs, for example, for metal surfaces at microwave frequencies. In this case the surface acts as a 'short circuit' to the incoming wave with zero transmitted wave, and hence the boundary condition on the tangential electric field requires that the total tangential electric field at the surface is zero. For this to occur the reflected wave must be 180° out of phase with the incident wave, and hence the reflection coefficient is -1. This phase shift on reflection should be embedded in the Fresnel equations, and it is interesting to ask how such phase shifts on reflection vary with dielectric constant, polarisation and angle of incidence.

To investigate this, we show in Figure 3.4 an example of the Fresnel equations evaluated for $\varepsilon_1 = 1$, $\varepsilon_2 = 4$. We make the following general observations from this example:

• The magnitude of the HH channel is always greater than or equal to VV. In fact, |VV| decreases with increasing angle of incidence and reaches zero at the Brewster angle (for complex dielectric constants the Brewster angle reaches a minimum of reflection but not exactly zero). This angle is a function of the ratio of dielectric constants of the two materials, as we now show. By considering the reflection process as scattering by electric dipoles (see equation (1.10)) in the lower medium, we can obtain a zero reflected signal under condition that the sum of refracted and reflected angles is $\pi/2$. In this case the reflected component lies along the null



Fig. 3.4 Fresnel reflection and transmission coefficients for $\varepsilon_r = 4$

in the dipole radiation pattern. Using Snell's law we can then obtain an expression for this Brewster angle θ_B , as shown in equation (3.18):

$$\theta_r + \theta_t = \frac{\pi}{2} \Rightarrow n_1 \sin \theta_i = n_2 \cos \theta_i \Rightarrow \tan \theta_B = \sqrt{\frac{\varepsilon_{r2}}{\varepsilon_{r1}}}$$
 (3.18)

Finally, we note that at normal incidence ($\theta = 0^{\circ}$) both HH and VV have the same magnitude but a 180-degree phase difference (see below for a discussion of phase and coordinates), while at grazing incidence ($\theta = 90^{\circ}$) both |HH| and |VV| have unit magnitude and zero phase difference. Hence there must be a switch in phase for one of the polarisation channels at some angle of incidence.

• The phase of the HH reflected signal is always 180 degrees, for $\varepsilon_2 > \varepsilon_1$. From Figure 3.2 we see that this means that the perpendicular component (H) undergoes a 180-degree phase shift on reflection, as discussed in the limit of perfect conductors. This is a direct consequence of the boundary conditions and coordinate system employed. The VV channel phase has a more complicated structure. We first note that the VV channel has zero phase for angles of incidence less than the Brewster angle and a 180-degree phase shift thereafter. This can be explained by inspection of the coordinate systems employed in Figure 3.2, as follows. From Figure 3.2 we see that the 'zero' VV phase is actually a consequence of the same 180° phase change on surface reflection that is seen in the HH component (consider the case as θ_i tends to zero). The reason it appears as a zero phase is that the surface coordinate definition of in-phase for polarisation components parallel to the plane of incidence is that their z components are parallel.

To further illustrate this coordinate-dependent phase notation, we show in Figure 3.5 the coordinates used for a sensor-oriented (BSA) description of normal incidence reflection. Note that by using the same logic as above, the HH reflection coefficient will now be positive (with the 180-degree surface phase shift now matching the coordinate shift of the BSA system), while VV



Fig. 3.5 Sensor (BSA) coordinate system for describing normal incidence Fresnel

reflection

will be negative. Hence we obtain the modified normal incidence reflection matrix based on the Fresnel equations as shown in equation (3.19) (see also equation (1.152)), which was used to obtain the specular backscatter matrix in equation (1.153).

$$[R]_{normal} = \begin{bmatrix} -R_{HH} & 0\\ 0 & -R_{VV} \end{bmatrix} = \frac{n_2 - 1}{n_2 + 1} \begin{bmatrix} 1 & 0\\ 0 & -1 \end{bmatrix}$$
(3.19)

Now, however, we are in a position to generalize this result to arbitrary angle of incidence. The sensor coordinate scattering matrix for Fresnel reflection then has the general form shown in equation (3.20):

$$[S]_{sensor} = A \begin{bmatrix} S_{HH} & 0\\ 0 & S_{VV} \end{bmatrix} \Rightarrow \begin{cases} S_{HH} = \frac{\sqrt{\frac{\varepsilon_2}{\varepsilon_1} - \sin^2 \theta_i} - \cos \theta_i}{\cos \theta_i + \sqrt{\frac{\varepsilon_2}{\varepsilon_1} - \sin^2 \theta_i}} \\ S_{VV} = \frac{\frac{\varepsilon_2}{\varepsilon_1} \cos \theta_i - \sqrt{\frac{\varepsilon_2}{\varepsilon_1} - \sin^2 \theta_i}}{\frac{\varepsilon_2}{\varepsilon_1} \cos \theta_i + \sqrt{\frac{\varepsilon_2}{\varepsilon_1} - \sin^2 \theta_i}} \end{cases}$$
(3.20)

with the important special case of normal incidence ($\theta = 0^{\circ}$) yielding a scalar multiple of the identity matrix, as shown in equation (3.21):

$$[S]_{BSA} = A \begin{bmatrix} 1 & 0\\ 0 & 1 \end{bmatrix}$$
(3.21)

3.1.2 Polarisation properties of surface backscatter

We now consider the important special case of surface backscatter. We begin with the Fresnel equations and note that, aside from 'specular reflection' at normal incidence, the energy is all reflected away from the backscatter direction. Hence, perfectly smooth surfaces have zero backscatter for oblique incidence. However, most natural surfaces are not smooth, and so some backscatter signal is observed. In this section we develop the polarisation properties of the backscattering matrix for such rough surfaces.

The first modification we must make to the Fresnel equations is to allow for the finite size of any surface element or 'facet'. Figure 3.6 shows a schematic representation of such an elementary scatterer. As with all scattering problems the procedure is to estimate the currents induced on the surface, and then use them as sources in the wave equation to estimate the scattering matrix. One of the simplest methods is to employ the Physical Optics (PO) current approximation (Ishimaru, 1991; Jones, 1989). This states that the current induced on the facet is the same as that induced on an infinite local tangent plane. The only change from the Fresnel equations is then to account for the finite extent L of the scatterer. This involves evaluation of an integral over the surface similar to that employed in the Huygens source of Figure 1.5. This result again yields a Fourier transform relationship between the current distribution and the scattered field. In this example we have a simple uniform current of limited physical extent L, which by Fourier transform yields a SINC function for the scattered field, so that in this case the BSA scattering matrix is a simple modification of the



Fig. 3.6 Backscatter from a square surface 'facet' of dimension L

Fresnel equations for an infinite plane, as shown in equation (3.22):

$$\begin{bmatrix} S(\theta \quad \lambda) \end{bmatrix} = \begin{bmatrix} S_{\perp} & 0\\ 0 & S_{\parallel} \end{bmatrix}_{BSA}$$

$$= \frac{2\sqrt{\pi L}}{\lambda} \cos \theta \frac{\sin (\beta L \sin \theta)}{\beta L \sin \theta} \begin{bmatrix} S_{HH}(\theta \quad \varepsilon_r) & 0\\ 0 & S_{VV}(\theta \quad \varepsilon_r) \end{bmatrix}$$

$$\Rightarrow \begin{cases} S_{HH} = \frac{\sqrt{\frac{\varepsilon_2}{\varepsilon_1} - \sin^2 \theta}}{\cos \theta + \sqrt{\frac{\varepsilon_2}{\varepsilon_1} - \sin^2 \theta}} \\ S_{VV} = \frac{\frac{\varepsilon_2}{\varepsilon_1} \cos \theta - \sqrt{\frac{\varepsilon_2}{\varepsilon_1} - \sin^2 \theta}}{\frac{\varepsilon_2}{\varepsilon_1} \cos \theta + \sqrt{\frac{\varepsilon_2}{\varepsilon_1} - \sin^2 \theta}} \end{cases}$$
(3.22)

This equation shows that, for a *finite* sized facet, we will always receive some backscattered signal, and that the polarisation of the backscattered wave is given by the Fresnel equations at the local angle of incidence. Note that if the surface facet is large compared to a wavelength ($\beta L >> 1$) then the SINC function falls quickly to zero with increasing θ , and so the backscattered amplitude is very small at all but normal incidence. This is called a specular surface return, and has a straightforward polarisation behaviour described by a multiple of the identity matrix, as shown in equation (3.21). Such approximations are at the heart of high-frequency rough surface scattering theory based on stationary phase evaluation of integrals (Jones, 1989; Ishimaru, 1991). In these cases it is no surprise, therefore, that in the absence of multiple scattering it is found that HH = VV and HV = 0. This is trivial polarisation behaviour (although the actual evaluation of the integrals can be involved), and is not typical of the microwave properties of natural surfaces (see Chapter 9).



$$[R] = [R_B] \cdot [R_A] = \begin{bmatrix} R_{HHB}R_{HHA} & 0\\ 0 & R_{VVB}R_{VVA} \end{bmatrix}$$

$$R_{HHA} = \frac{\cos\theta - \sqrt{\varepsilon_A - \sin^2\theta}}{\cos\theta + \sqrt{\varepsilon_A - \sin^2\theta}} \quad R_{HHB} = \frac{\sin\theta - \sqrt{\varepsilon_B - \cos^2\theta}}{\sin\theta + \sqrt{\varepsilon_B - \cos^2\theta}}$$

$$R_{VVA} = \frac{\varepsilon_A \cos\theta - \sqrt{\varepsilon_A - \sin^2\theta}}{\varepsilon_A \cos\theta + \sqrt{\varepsilon_A - \sin^2\theta}} \quad R_{VVB} = \frac{\varepsilon_A \sin\theta - \sqrt{\varepsilon_A - \cos^2\theta}}{\varepsilon_A \sin\theta + \sqrt{\varepsilon_A - \cos^2\theta}}$$
(3.23)



Fig. 3.7 Retro-reflection or dihedral scattering geometry



Fig. 3.8 Normalized BSA scattering matrix elements for dihedral scattering for $\varepsilon_r = 25$ and for $\varepsilon_r = 4$

Making the usual coordinate corrections, the scattering matrix in the sensor (BSA) system now has the form shown in equation (3.24):

$$[S]_{BSA} = \begin{bmatrix} S_{\perp} & 0\\ 0 & S_{\parallel} \end{bmatrix} = A \begin{bmatrix} R_{HHB}R_{HHB} & 0\\ 0 & -R_{VVB}R_{VVA} \end{bmatrix}$$
(3.24)

The most important consequence of this is a 180-degree phase shift between HH and VV for a range of incidence angles centred on 45 degrees when compared to the single reflection matrices. Figure 3.8 shows an example for the HH and VV components of [S] for dihedral retroreflection at varying angles of incidence for $\varepsilon_A = \varepsilon_B = 2.25$ and $\varepsilon_A = \varepsilon_B = 9$. Notice how, as the dielectric constant increases so the width of the π phase shift zone increases. In the limit of a perfect conductor when $\varepsilon_r \rightarrow \infty$, such as metal surfaces at microwave frequencies, the scattering matrix for a dihedral retroreflector has a constant form for all angles of incidence, as shown in equation (3.25), which we used in equation (1.157).

$$[S]_{BSA} = A \begin{bmatrix} 1 & 0\\ 0 & -1 \end{bmatrix}$$
(3.25)

We see now that the polarimetric phase behaviour becomes more complicated when we start to consider scattering by dielectric media, and this will motivate the development of the 'alpha parameter' description of scattering in Chapter 4. First, however, we turn to another important surface scattering model, where polarisation behaviour can be fully quantified.

3.1.3 The Bragg surface scattering model

A second important example of surface scattering is when the surface facet is large compared to a wavelength but the surface is not smooth over the length scale L. In this case the solution for the scattered field can be obtained as a perturbation of that from the underlying smooth surface, and an analytical solution obtained for the scattering matrix in terms of an infinite series. This is termed the small perturbation method (SPM) (Tsang, 1985). Keeping only the first term in this series gives a good approximation to the solution when the perturbation



Fig. 3.9 Small perturbation or Bragg rough surface scattering and definition of the rms roughness parameter for an *N*-point discretisation of the surface

is limited, so that the rms roughness of the surface s is small compared to the wavelength $\beta = 2\pi/\lambda$; that is, β s is a small quantity. In practice, β s < 0.3 is a suitable condition, and so this constitutes a good low-frequency approximation, typically valid at L (1.2 GHz) and P (450 MHz) bands for remote sensing of natural surfaces. Figure 3.9 shows a schematic representation of this small-scale roughness. This is also called the vector small perturbation or Bragg scattering model (after William Lawrence Bragg (1890–1971), who first formulated the boundary condition required for such scattering). From our point of view the most important observation is that the perturbed boundary conditions lead to *different* polarisation coefficients from those obtained for the Fresnel case. The scattered field from an arbitrary rough surface characterized by a height function z(x,y), illuminated by a wavelength long enough for this small perturbation assumption to apply, is given as shown in equation (3.26) (Ulaby, 1986):

$$E_{pq}^{s} = i \, 2\beta \cos\theta \, B_{pq} \, \hat{Z} \, (\beta_{X} + \beta \sin\theta, \beta_{Y}) \\ \begin{cases} B_{HH} = \frac{(1 - \varepsilon_{r}) \cos\phi_{s}}{\left(\cos\theta_{s} + \sqrt{\varepsilon_{r} - \sin^{2}\theta_{s}}\right) \left(\cos\theta + \sqrt{\varepsilon_{r} - \sin^{2}\theta}\right)} \\ B_{VV} = \frac{(1 - \varepsilon_{r}) \left(\varepsilon_{r} \sin\theta \sin\theta_{s} - \sqrt{\varepsilon_{r} - \sin^{2}\theta_{s}} \sqrt{\varepsilon_{r} - \sin^{2}\theta} \cos\phi_{s}\right)}{\left(\varepsilon_{r} \cos\theta_{s} + \sqrt{\varepsilon_{r} - \sin^{2}\theta_{s}}\right) \left(\varepsilon_{r} \cos\theta + \sqrt{\varepsilon_{r} - \sin^{2}\theta}\right)} \\ B_{HV} = \frac{-(1 - \varepsilon_{r}) \sqrt{\varepsilon_{r} - \sin^{2}\theta} \sin\phi_{s}}{\left(\cos\theta_{s} + \sqrt{\varepsilon_{r} - \sin^{2}\theta_{s}}\right) \left(\varepsilon_{r} \cos\theta + \sqrt{\varepsilon_{r} - \sin^{2}\theta}\right)} \\ B_{VH} = \frac{(1 - \varepsilon_{r}) \sqrt{\varepsilon_{r} - \sin^{2}\theta_{s}} \sin\phi_{s}}{\left(\varepsilon_{r} \cos\theta_{s} + \sqrt{\varepsilon_{r} - \sin^{2}\theta_{s}}\right) \left(\cos\theta + \sqrt{\varepsilon_{r} - \sin^{2}\theta}\right)} \\ \end{cases}$$
(3.26)

where

$$\hat{Z}(\beta_x, \beta_y) = \frac{1}{2\pi} \iint_{\infty} z(x, y) e^{-i(\beta_x x + \beta_y y)} dx dy$$

$$\beta_x = -\beta \sin \theta_s \cos \phi_s$$

$$\beta_y = -\beta \sin \theta_s \sin \phi_s$$

$$\beta = \frac{2\pi}{\lambda}$$

(3.27)

For the special case of backscatter ($\theta_s = \theta$ and $\phi_s = \pi$) this yields the following important simplification:

1

$$B_{pq}^{s} = i \ 2\beta \cos\theta \ B_{pq} \ \hat{Z} \ (2\beta \sin\theta)$$

$$\Rightarrow \begin{cases} B_{HH} = B_{\perp} = \frac{\cos\theta - \sqrt{\varepsilon_r - \sin^2\theta}}{\cos\theta + \sqrt{\varepsilon_r - \sin^2\theta}} = R_{HH} \\ B_{VV} = B_{||} = \frac{(\varepsilon_r - 1) \left[\sin^2\theta - \varepsilon_r (1 + \sin^2\theta) \right]}{\left(\varepsilon_r \cos\theta + \sqrt{\varepsilon_r - \sin^2\theta} \right)^2} \neq R_{VV} \end{cases}$$

$$(3.28)$$

$$B_{HV} = B_{VH} = 0$$

where \hat{Z} is the Fourier transform of the surface profile. For backscattered power this becomes the Fourier transform of the surface correlation coefficient or the normalized roughness spectrum. Hence, to first order the backscatter depends only a particular frequency component of the surface roughness spectrum (Bragg scattering). But what of the polarisation properties of such scattering? To proceed we make the following observations.

- The HH component equals the HH Fresnel equation, but the VV component is very different from that predicted for Fresnel reflection. In particular there is no longer a Brewster angle effect.
- The BSA scattering matrix for Bragg surface backscatter then has the following form:

$$[S]_{BSA} = \begin{bmatrix} S_{\perp} & 0\\ 0 & S_{||} \end{bmatrix} = -i2\beta\cos\theta \,\hat{Z} \begin{bmatrix} B_{HH} & 0\\ 0 & B_{VV} \end{bmatrix}$$
(3.29)

Figure 3.10 shows the variation of the polarisation coefficients $B_{\rm HH}$ and $B_{\rm VV}$ with angle of incidence for backscatter from a surface with $\varepsilon_{\rm r} = 4$. Compare the results with those for Fresnel reflection in Figure 3.4. Note that in this case the *B* coefficients are *not* reflection coefficients. The actual level of backscattered



Fig. 3.10 Bragg coefficients for backscatter from a rough surface with for $\varepsilon_r = 4$

signal depends on the component of the Fourier transform of the surface (satisfying the Bragg condition). Note especially the absence of a Brewster angle in the VV channel, and also note that the magnitude of VV is always greater than or equal to HH—the opposite trend to that observed for Fresnel reflection.

As a final observation we note that the effect of surface roughness is contained in the Fourier transform component \hat{Z} , which is common to all polarisation channels. Therefore, if we take a ratio of scattered signals, such as HH/VV, then the effects of roughness will cancel, and the ratio depends only on the angle of incidence and dielectric constant of the surface. This provides an important stimulus to using polarisation diversity for surface parameter retrieval in remote sensing. For example, we note that as the angle of incidence increases so the difference between HH and VV increases, and that this ratio is a function of dielectric constant and hence an indicator of soil moisture or salinity. One way to represent these variations is by using the scattering alpha parameter, discussed in Chapter 4. According to this representation we can represent the scattering ratio in a rotation invariant manner by forming the following function:

$$\alpha_b = \tan^{-1} \left(\left| \frac{B_{HH} - B_{VV}}{B_{HH} + B_{VV}} \right| \right) \quad 0 \le \alpha \le \frac{\pi}{2}$$
(3.30)

Figure 3.11 shows how alpha changes with angle of incidence for three cases of dielectric constant, $\varepsilon_r = 2$, $\varepsilon_r = 9$, and $\varepsilon_r = 81$. We note that the dynamic range in alpha available to distinguish between wet and dry surfaces increases with angle of incidence. By taking the difference between the curves we obtain an estimate of this dynamic range, as shown in Figure 3.12. This can be used, for example, to assess the polarimetric calibration requirements of a sensor in order to be able to distinguish wet and dry surfaces. We note that for angles of incidence less than 25 degrees (typical of spaceborne radar systems) the available dynamic range is limited to less than 5 degrees of alpha variation.



Fig. 3.11 Variation of Bragg alpha angle with angle of incidence for three dielectric constants



Fig. 3.12 Dynamic range of alpha between low and high dielectrics as a function of angle of incidence



3.1.4 Coherent surface scattering component

Large-scale surface roughness also has an effect on the dihedral scattering mechanism, introduced in Figure 3.7, by reducing the level of specular or coherent reflection from a surface at the expense of a diffuse or noncoherent component (Ulaby, 1982). Hence the Fresnel coefficients are modified by an attenuating multiplicative factor F that depends on surface roughness and angle of incidence (of the form shown in equation (3.31)). Importantly for our purposes, such a factor is independent of polarisation, and therefore does not change the scattering alpha parameter. Both surfaces involved in the dihedral return will in general be effected by such factors, but it is sufficient for us to consider a single such reflection as shown schematically in Figure 3.13. The modified form of reflection coefficient in the presence of rough surface scattering is then shown in equation (3.31):

$$F = e^{-2\beta s \cos\theta} \rightarrow [R] = F \begin{bmatrix} \frac{\cos\theta - \sqrt{\varepsilon_A - \sin^2\theta}}{\cos\theta + \sqrt{\varepsilon_A - \sin^2\theta}} & 0\\ 0 & \frac{\varepsilon_A \cos\theta - \sqrt{\varepsilon_A - \sin^2\theta}}{\varepsilon_A \cos\theta + \sqrt{\varepsilon_A - \sin^2\theta}} \end{bmatrix}$$
(3.31)

Here s is the surface rms roughness and β the free space wavenumber $2\pi/\lambda$. Figure 3.14 illustrates how the power attenuation factor varies for typical roughness variations (2, 4 and 6 cm rms) at L-band frequency (1.3 GHz). There are



Fig. 3.14 Example attenuation of surface specular reflection due to surface roughness factor F $\,$

Fig. 3.15 Dihedral alpha parameter for a 22.5-degree angle of incidence and dielectric variations in the two surfaces

three important consequences of this result for polarimetric studies. The first is that in the presence of roughness the dihedral return, even though it is based on a strong specular scattering effect, can be attenuated to a level where the direct surface return can be dominant. Thus in practice we must consider the possibility that both phenomena may be present at the same time. This will lead us to develop so-called decomposition theorems for the interpretation of radar scattering (see Chapter 4).

The second key idea is that any ratio of elements of the scattering matrix is unchanged by the presence of surface roughness. As in the case of Bragg scattering, we can then calculate the scattering alpha parameter (from equation (3.30)) as a function of angle of incidence and dielectric constants of the two surfaces involved. Figures 3.15 and 3.16 show examples of the dynamic range of alpha for a shallow angle $\theta = 22.5^{\circ}$ and a steeper angle $\theta = 45^{\circ}$. We note the following key points:



Fig. 3.16 Dihedral alpha parameter for a 45-degree angle of incidence and dielectric variations in the two surfaces

- 1. The shallow angle alpha shows little dependence on the first dielectric component (from the horizontal surface) but is directly related to the second (from the vertical), due to its effective higher angle of incidence.
- 2. The steeper angle alpha shows more dependence on *both* dielectric constants (horizontal and vertical surface reflections) and hence, for any given alpha value, we can only ever obtain an estimate of the product of the two dielectric constants.
- 3. The dynamic range in alpha is much larger, even for the shallow angle, than that obtained for direct Bragg backscattering (Figure 3.12).
- 4. The alpha angles are nearly always greater than $\pi/4$, as expected for dihedral scattering, but can go lower than this for very dry materials and angles of incidence where the Brewster angle effect causes a switch in the scattered phase.

Finally, we note that one other advantage of the alpha parameter formulation is its invariance to rotations in the plane of polarisation. Such rotations can occur, for example, when we consider scattering from sloped surfaces. In such cases, although the alpha parameter can remain unchanged, crosspolarisation is introduced to the scattering vector. We now turn to consider slope effects on the polarisation properties of surface scattered waves.

3.1.5 The effect of surface slope on the scattering matrix

In practice the surface facet normal may have an arbitrary orientation in space due to surface slope or large-scale roughness. If we establish a right-handed coordinate system centred on the facet then we can write the general facet normal vector as shown in Figure 3.17. We can further define two principal slope components from this normal (Lee, 2000; Schuler, 2002). Rotation about the x axis we then call range slope, defined by an angle γ . Similarly, rotation about the y axis we call azimuth slope ψ . These slopes are direct parameters of interest for estimation in radar remote sensing, and so we prefer to use them rather than <u>n</u> in the following equations. A wave is now incident in the zy plane


Fig. 3.17 Arbitrary surface normal orientation and range and azimuth slopes

at an angle $\bar{\phi}$ to the z axis. Hence if $\underline{n} = (0, 0, 1)$ then we recover the situation already treated above. However, in general this will not always be the case, and we must therefore account for the effects of arbitrary local surface normal on the scattering amplitude matrix for the facet. The key extension we now require is to define a local tangent vector \underline{t} defined in terms of the incident wave vector β and the normal \underline{n} , as shown in equation (3.32):

$$\underline{t} = \frac{\underline{n} \times \underline{\beta}}{\left|\underline{n} \times \underline{\beta}\right|} \tag{3.32}$$

With this local coordinate system in place we must now modify the scattering matrix in two ways:

• The angle of incidence θ used for evaluation of the Bragg or Fresnel coefficients is no longer simply $\bar{\phi}$ in Figure 3.17, but is now defined from the inner product between the surface normal and incident wave vector, as shown in equation (3.33):

$$\cos\theta = \underline{n}.\underline{\beta} = -n_2\sin\bar{\phi} + n_3\cos\bar{\phi} = \frac{\tan\gamma\sin\phi + \cos\phi}{\sqrt{1 + \tan^2\gamma + \tan^2\psi}} \quad (3.33)$$

• The combined effect of range and azimuth slopes causes an effective rotation of the surface in the plane of polarisation through an angle χ , as shown in equation (3.34):

$$\frac{\cos \chi = \underline{h}.\underline{t}}{\sin \chi = -|\underline{h} \times \underline{t}|} \Rightarrow \tan \chi = \frac{-n_1}{n_2 \cos \overline{\phi} + n_3 \sin \overline{\phi}} \\
= \frac{\tan \psi}{\sin \overline{\phi} - \cos \overline{\phi} \tan \gamma}$$
(3.34)

The combined effect of these two angles is to modify the scattering matrix of the facet to that shown in equation (3.35), which we see leads to the generation of crosspolarisation from the facet:

$$\begin{bmatrix} S(\underline{n}) \end{bmatrix} = \begin{bmatrix} \cos \chi & \sin \chi \\ -\sin \chi & \cos \chi \end{bmatrix} \cdot \begin{bmatrix} S_{\perp}(\theta) & 0 \\ 0 & S_{||}(\theta) \end{bmatrix} \cdot \begin{bmatrix} \cos \chi & -\sin \chi \\ \sin \chi & \cos \chi \end{bmatrix}$$
(3.35)

Indeed, by measuring the level of crosspolarisation we can in principle estimate χ by employing an SVD of the scattering matrix and obtain remote information about the slope of the facet. Slope also modifies the apparent alpha parameter for the surface backscatter. The range slope component will cause an increase (for slopes away from the radar) or decrease (for slopes towards the radar) in the apparent alpha.

Finally, we note the effects of slope on the specular dihedral response (see Figure 3.7). The key consequence of slope in this context is that the backscatter ray path now no longer includes the two specular reflection mechanisms for surfaces A and B (since the angle between the two normals is no longer $\pi/2$). Hence the strongest return is now scattered into a small bistatic angle (given by the slope) and is not returned in the exact backscatter direction. For this reason the presence of slopes can attenuate the backscatter dihedral return (especially when combined with roughness, as mentioned in equation (3.31)) and leave only the direct surface component. The slope tolerance of the specular dihedral component depends on many factors, such as the height of the vertical scatterer and the radar wavelength. Instead of considering this phenomenon on a caseby-case basis we instead prefer, in polarimetry, to model 'surfaces' as some a priori unknown mixture of direct and specular dihedral scattering mechanisms, and attempt to retrieve the properties and relative amplitudes of these from the data itself. These are treated in Chapter 4 under the general topic of decomposition theorems. First, however, we consider the topic of depolarisation by surfaces.

3.2 Surface depolarisation

We have seen that surface scattering provides a diversity of polarisation responses, varying with angle of incidence, dielectric constant and surface slope. In particular, we have seen that slope leads to crosspolarisation in surface scattering; but this is a deterministic phenomenon and can be removed, for example, by a singular value decomposition of the scattering matrix. However, so far we have ignored any reference to depolarisation. This can and does arise in natural surface scattering, and is difficult to quantify in the context of rigorous scattering models (Borgeaud, 1994; Hajnsek, 2003). In this section we consider one simple low-frequency analytic model that has been developed to include full wave depolarisation: the extended or X-Bragg model (Hajnsek, 2003; Allain, 2003).

3.2.1 The extended or X-Bragg model

Rough surface scattering provides an important example of reflection symmetric depolarisation, where the mean surface normal provides an obvious axis of symmetry. In the limit of a smooth surface $\beta s < 0.3$, where β is the wavenumber and s is the rms roughness, the small perturbation model (SPM) or Bragg scattering applies (see equation (3.29)). In this case the backscattered electric field is related to the component of the Fourier transform of the surface profile 'resonant' with the incident wave. The scattering coefficients are determined by the boundary conditions and vary with polarisation, as shown in equation (3.36). A key consequence of this model is that the effects of surface roughness act as

a scalar multiplier in all polarisation channels, and so all ratios of polarisation remain independent of roughness and depend only on dielectric constant and angle of incidence through the complex $B_{\rm HH}$ and $B_{\rm VV}$ coefficients. Such a surface therefore has zero depolarisation (although possibly crosspolarisation if the normal is misaligned with the wave coordinates). This is seen in equation (3.36), where [*T*] has only one non-zero eigenvalue and the scattering entropy H is zero. The surface dielectric properties can then be represented by the scattering mechanism α , as shown in equation (3.30). However, smooth natural surfaces are observed to depolarise incident waves, and so this model needs some modification. To achieve this we first go to the other extreme of a very rough surface when $\beta s >> 1$. In this case we propose that the surface acts as an azimuthally symmetric depolariser with a coherency matrix of the general form shown in equation (3.37). This has scattering entropy greater than zero, and hence causes depolarisation of the incident wave.

$$E_{pq}^{s} = i \, 2\beta \cos\theta \, B_{pq} \, \hat{Z} \, (2\beta \sin\theta)$$

$$\Rightarrow \begin{cases} B_{HH} = B_{\perp} = \frac{\cos\theta - \sqrt{\varepsilon_{r} - \sin^{2}\theta}}{\cos\theta - +\sqrt{\varepsilon_{r} - \sin^{2}\theta}} \\ B_{VV} = B_{||} = \frac{(\varepsilon_{r} - 1) \left[\sin^{2}\theta - \varepsilon_{r}(1 + \sin^{2}\theta)\right]}{\left(\varepsilon_{r} \cos\theta + \sqrt{\varepsilon_{r} - \sin^{2}\theta}\right)^{2}} \\ B_{HV} = B_{VH} = 0 \end{cases}$$
(3.36)

$$\stackrel{\beta_{S \ll 1}}{\longrightarrow} [T_{S}] = m_{s} \begin{bmatrix} a & b & 0 \\ b^{*} & c & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

$$= \begin{bmatrix} \underline{u}_{1} & \underline{u}_{2} & \underline{u}_{3} \end{bmatrix} \cdot \begin{bmatrix} \lambda_{1} & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \cdot \begin{bmatrix} \underline{u}_{1} & \underline{u}_{2} & \underline{u}_{3} \end{bmatrix}^{*T}$$

$$\Rightarrow H_{s} = 0$$

We now develop a model of a depolariser based on a smooth transition between these two limits, reducing to equation (3.36) for smooth surfaces and to equation (3.37) for rough surfaces.

$$\stackrel{\beta s >>1}{\longrightarrow} [T_s] = t_{11} \begin{bmatrix} 1 & 0 & 0\\ 0 & m & 0\\ 0 & 0 & m \end{bmatrix} \Rightarrow H_s > 0$$
 (3.37)

One way to do this is to assume that the major perturbation to equation (3.36) arises from micro-variations in surface slope as roughness increases. There is a single parameter in the scattering vector that is influenced strongly by changes in slope: χ , as shown in equation (3.38) (see also equation (3.35)). In this case we can therefore propose that depolarisation is primarily caused by integration over a distribution of χ . Even if the exact details of this distribution (p(χ)) are

unknown, the consequences for the form of the resulting coherency matrix are important, as shown in equation (3.38):

$$\begin{bmatrix} T_{S} \end{bmatrix} = m_{s} \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos 2\chi & \sin 2\chi \\ 0 & -\sin 2\chi & \cos 2\chi \end{bmatrix} \begin{bmatrix} a & b & 0 \\ b^{*} & c & 0 \\ 0 & 0 & 0 \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos 2\chi & -\sin 2\chi \\ 0 & \sin 2\chi & \cos 2\chi \end{bmatrix}$$
$$\tau = \int \cos 2\chi \cdot p(\chi) \, d\chi$$
$$\delta = \int \cos^{2} 2\chi \cdot p(\chi) \, d\chi \end{bmatrix} \rightarrow [T_{S}] = m_{s} \begin{bmatrix} a & b\tau & 0 \\ b^{*}\tau & c\delta & 0 \\ 0 & 0 & c(1-\delta) \end{bmatrix} \Rightarrow H_{s} > 0$$
(3.38)

In the general case we see we must introduce two depolarising parameters δ and τ , both of which are related to the (as yet unknown) distribution $p(\chi)$. However, we see that this model has the correct boundary conditions: when $\delta = 1, \tau = 1$ (smooth surfaces), so this model tends to the zero entropy Bragg model (equation (3.36)), while when τ tends to zero and δ tends to 0.5 (rough surfaces), so the surface tends to an azimuthally symmetric depolariser (equation (3.37)). Significantly, between these two extremes we note that the contribution of surface roughness and dielectric constant properties to depolarisation can be separated (regardless of the distribution $p(\chi)$) by using the two secondary parameters R and M, shown in equations (3.39) and (3.40) (Cloude, 2002a):

$$R(\delta) = \frac{t_{22} - t_{33}}{t_{22} + t_{33}} = \frac{\langle |S_{HH} - S_{VV}|^2 \rangle - 4 \langle |S_{HV}|^2 \rangle}{\langle |S_{HH} - S_{VV}|^2 \rangle + 4 \langle |S_{HV}|^2 \rangle} = 2\delta - 1 \quad 0 \le R \le 1$$
(2.20)

21

$$M(\theta, \varepsilon_r) = \frac{t_{22} + t_{33}}{t_{11}} = \frac{\langle |S_{HH} - S_{VV}|^2 \rangle + 4 \langle |S_{HV}|^2 \rangle}{\langle |S_{HH} + S_{VV}|^2 \rangle} = \tan^2 \alpha_b$$
(3.40)

,

In this way R can be considered a roughness measure of the surface based on its level of depolarisation—R = 0 corresponding to a very rough surface, and R = 1 to very smooth. Such a parameter is also insensitive to variations in surface dielectric constant, and hence is robust to changes in surface moisture content and salinity, for example. On the other hand, M is a material indicator, increasing with increasing dielectric constant of the surface while remaining insensitive to roughness variations. Note that M also depends on angle of incidence θ . As θ increases from zero (normal incidence), so the variation of M with dielectric constant increases. R, on the other hand, is independent of θ , and therefore does not vary with either dielectric constant or changes in angle of incidence.

The roughness factor *R* behaves like a coherence amplitude, lying between 0 and 1. In fact, R equals the coherence between the LL and RR circular polarisations, as shown in equation (3.41). This parameter was first proposed and developed as a surface discriminator in Mattia (1997).

$$\gamma_{LLRR} = \frac{\langle s_{LL} s_{RR}^* \rangle}{\sqrt{\langle s_{LL} s_{LL}^* \rangle \langle s_{RR} s_{RR}^* \rangle}}$$

$$= \frac{\langle (p_1 + ip_2) (-p_1 + ip_2)^* \rangle}{\sqrt{\langle (p_1 + ip_2) (p_1 + ip_2)^* \rangle \langle (-p_1 + ip_2) (-p_1 + ip_2)^* \rangle}}$$

$$\begin{cases} p_1 = S_{HH} - S_{VV} \\ p_2 = 2S_{HV} \end{cases}$$

$$\xrightarrow{\text{reflction symmetry}} \gamma_{LLRR} = \frac{\langle |p_2|^2 \rangle - \langle |p_1|^2 \rangle}{\langle |p_1|^2 \rangle + \langle |p_2|^2 \rangle} = -R$$

$$(3.41)$$

Note that in the presence of a surface with mean azimuth slope $\overline{\chi}$, the X-Bragg coherency matrix has the modified form shown in equation (3.42):

$$[T_{\mathcal{S}}] = m_{\mathcal{S}} \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos 2\overline{\chi} & \sin 2\overline{\chi} \\ 0 & -\sin 2\overline{\chi} & \cos 2\overline{\chi} \end{bmatrix} \begin{bmatrix} a & b\tau & 0 \\ b^{*}\tau & c\delta & 0 \\ 0 & 0 & c(1-\delta) \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos 2\overline{\chi} & -\sin 2\overline{\chi} \\ 0 & \sin 2\overline{\chi} & \cos 2\overline{\chi} \end{bmatrix}$$
(3.42)

from which it follows that the revised diagonal elements of [T] have the form shown in equation (3.43):

$$t_{11} = a$$

$$t_{22} = c(\delta \cos 4\overline{\chi} + \sin^2 2\overline{\chi}) \qquad (3.43)$$

$$t_{33} = c(\cos^2 2\overline{\chi} - \delta \cos 4\overline{\chi})$$

which implies that while the material index M remains invariant to azimuth slope, the roughness indicator R varies as $\cos 4\overline{\chi}$; therefore, as mean slope increases so the apparent roughness increases. We note, however, that by estimating the LLRR coherence in both amplitude and phase we can compensate for such slope effects ($\overline{\chi}$ from the mean phase of the coherence, see Chapter 4) and obtain a basis invariant estimate of the roughness (R from the coherence amplitude, corrected for χ). In this way we see that circular polarisations and their coherence are of great utility in surface scattering studies.

As this approach is based on the low-frequency Bragg model but provides a link to high-frequency depolarisation effects it is termed the extended or X-Bragg surface scattering model. We can formulate this model for a geometrical representation of its depolarisation properties in the backscatter entropy/alpha plane as follows. If we assume for simplicity a uniform distribution of slopes for $p(\chi)$ with width Δ , then the X-Bragg model takes on the explicit form shown in equation (3.44), where the coefficients *a*, *b* and *c* are given in terms of the Bragg scattering coefficients:

$$\tau = \int \cos 2\chi p(\chi) d\chi = \frac{\sin 2\Delta}{2\Delta}$$

$$\delta = \int \cos^2 2\chi p(\chi) d\chi = \frac{1}{2} \left(1 + \frac{\sin 4\Delta}{4\Delta} \right)$$

$$\to [T_S] = m_s \begin{bmatrix} a & b \frac{\sin 2\Delta}{2\Delta} & 0 \\ b^* \frac{\sin 2\Delta}{2\Delta} & c \frac{1}{2} \left(1 + \frac{\sin 4\Delta}{4\Delta} \right) & 0 \\ 0 & 0 & c \frac{1}{2} \left(1 - \frac{\sin 4\Delta}{4\Delta} \right) \end{bmatrix}$$
(3.44)

$$a = |B_{HH} + B_{VV}|^{2}$$

$$b = (B_{HH} + B_{VV})(B_{HH} - B_{VV})^{*} \Rightarrow \begin{cases} B_{HH} = \frac{\cos\theta - \sqrt{\varepsilon_{r} - \sin^{2}\theta}}{\cos\theta + \sqrt{\varepsilon_{r} - \sin^{2}\theta}} \\ B_{VV} = \frac{(\varepsilon_{r} - 1)\left[\sin^{2}\theta - \varepsilon_{r}(1 + \sin^{2}\theta)\right]}{\left(\varepsilon_{r}\cos\theta + \sqrt{\varepsilon_{r} - \sin^{2}\theta}\right)^{2}} \end{cases}$$

We can now consider the loci of H/α points generated when Δ varies from 0 to π for a fixed angle of incidence and varying dielectric constant ε_r . Such loci are shown in Figure 3.18, for θ varying from 20 to 50 degrees. We have chosen a high dielectric constant ($\varepsilon_r = 75$), so these loci represent the upper bounds on α for each angle of incidence. Again we note that for each angle of incidence the alpha angle changes only little with roughness changes, as implied by the *M* parameter in equation (3.40). We note that as the angle of incidence increases,



Fig. 3.18 Geometrical representation of the X-Bragg model in the entropy/alpha plane

so the potential depolarisation increases, and at small angles close to normal the loci lie closely packed around the origin. At normal incidence itself, the locus reduces to a point at the origin for all dielectric constants and roughness variations.

A second, related approach to modelling surface depolarisation was developed in Alain (2003). In this approach the covariance matrix for the surface backscatter is modelled in two components: a single scattering 'coherent' element (similar to Bragg), and a multiple scattering component, itself calculated from two surface parameters—the rms roughness and surface correlation length—using the integral equation model (IEM) first developed by Fung (1992). In this model, surface scattering depends not only on rms height but also on surface correlation length L_c . Two of the most popular correlation models for

surface studies are the Gaussian correlation function $(e^{-\frac{r^2}{L_c}})$ and the exponential $(e^{-\frac{|r|}{L_c}})$, where *r* is the separation of two points on the surface. This composite model is then summarized in equation (3.45):

$$[C] = [C]^{S} + [C]^{M} = \begin{bmatrix} c_{hhhh}^{s} & 0 & c_{hhvv}^{s} \\ 0 & 0 & 0 \\ c_{vvhh}^{s} & 0 & c_{vvvv}^{s} \end{bmatrix} + \begin{bmatrix} c_{hhhh}^{m} & 0 & c_{hhvv}^{m} \\ 0 & c_{hvhv}^{m} & 0 \\ c_{vvhh}^{m} & 0 & c_{vvvv}^{m} \end{bmatrix}$$
(3.45)

As multiple scattering becomes stronger, so the second term dominates (which is diagonally dominant), leading to an increase in entropy and depolarisation. Detailed comparisons between this model and X-Bragg show similar features in the entropy/alpha plane.

So far we have dealt only with surface backscatter (N = 3 depolarisation). We now turn to consider the treatment of the more general N = 4 bistatic surface scattering case.

3.2.2 Polarisation effects in bistatic surface scattering

The geometry of the general bistatic surface scattering problem is shown in Figure 3.19. An incident wave is scattered by a rough surface, and we are interested in the polarisation properties of radiation scattered into an arbitrary direction. The geometry of this problem is more complicated than for backscatter, there being three primary angles to consider: θ_i , the incident polar angle; θ_r , the reflected polar angle; and $\phi = \phi_r - \phi_r$, the difference in azimuthal angles for



Fig. 3.19 Geometry of bistatic surface scattering

incident and reflected waves. In particular we are often interested in so-called out-of-plane behaviour, when the surface normal, incident and scattered wave vectors do not all lie in a plane. A fully polarimetric characterization of this problem therefore involves knowledge of sixteen functions of three parameters (the elements of [M] or [T]). This set defines the polarimetric bidirectional reflectance distribution function (BRDF) $f(\theta_i, \theta_r, \phi)$, and its full characterization, either by measurement or theory, is a complicated process (Mendez, 1987; Priest, 2000). For this reason, various simplified physics-based models have been proposed in the literature. One of the most common of these is the microfacet model that we now consider (Priest, 2000).

First however, there are several important physical properties of the BRDF to be considered. One of these relates to the integration of the function over a hemisphere, to obtain the directional hemispherical reflectivity or DHR which provides a normalization condition, as shown in equation (3.46):

$$DHR(\theta_i) = \int f(\theta_i, \theta_r, \phi) \cos \theta_r d^2 \omega = \iiint f(\theta_i, \theta_r, \phi) \cos \theta_r \sin \theta_r d\theta_r d\phi$$
(3.46)

Note that the $\cos \theta_r$ factor appears because the BRDF is defined per-unit projected area. As a consequence we note that $f(\theta_i, \theta_r, \phi)$ is related to σ_s , the scattering cross-section per-unit illuminated area as shown in equation (3.47):

$$f = \frac{\sigma_s}{\cos \theta_i \cos \theta_r} \tag{3.47}$$

As an extreme case, a Lambertian surface is defined so that the outgoing power per unit solid angle, per-unit projected area is independent of angle. Hence $f = 1/\pi$ for a Lambertian surface; and it also acts as an ideal depolariser with scattering entropy H = 1, so that it has a coherency matrix of the form shown in equation (3.48). Natural surfaces, however, do not show this type of depolarising behaviour, and generally have entropy less than 1 (see Chapter 9).

$$\langle [T_{lambertian}] \rangle = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$
(3.48)

A second important property of the BRDF relates to the reciprocity theorem. By exchanging θ_r and θ_i in equation (3.46) and requiring invariance of the DHR leads to the reciprocity or exchange symmetry as shown in equation (3.49):

$$f(\theta_i, \theta_r, \phi) = f(\theta_r, \theta_i, \phi) \tag{3.49}$$

Having established these basic properties of the BRDF, we now turn to consider the microfacet model, the physics of which is rooted in geometrical optics. In its basic form it postulates scattering from a collection of randomly oriented microfacets (see Figure 3.6) comprising the rough surface. The geometrical distribution of these facets, together with a local reflection law—usually the Fresnel equations and Snell's law—then defines the structure of the model. One key feature of the microfacet model for our purposes is that it reduces the complexity of the angular dependencies to a familiar similarity transformation of the single facet scattering matrix, as we now demonstrate. The starting point is to convert the three primary scattering angles into four secondary geometrical parameters. The first is an angle Ω , which is the local angle of incidence (and angle of reflection) onto the facet, and is important in determining the physical properties of scattering by a facet. This angle is defined from the inner product of local surface normal unit vector <u>n</u> and direction vector of the incident wave <u>r</u>_i, as shown in equation (3.50):

$$\cos \Omega = \underline{n} \cdot \underline{r}_i \Rightarrow \cos 2\beta = \cos \theta_i \cos \theta_r + \sin \theta_r \sin \theta_r \cos (\phi_i - \phi_r) \quad (3.50)$$

From the local reflection law we can also relate Ω to the reflected wave vector \underline{r}_r , as shown in equation (3.51):

$$\underline{r}_i + \underline{r}_r = 2\cos\Omega\underline{n} \tag{3.51}$$

This leads us to the second important angle, θ , being the polar angle of the local normal, as shown in equation (3.52):

$$\cos\theta = \frac{\cos\theta_i + \cos\theta_r}{2\cos\Omega} \tag{3.52}$$

Each microfacet is then characterized by its normal \underline{n} , which is further assumed to be symmetrically distributed about the z axis according to some given probability function. A commonly used example is the Gaussian surface height distribution, which in normalized form can be written in terms of the local surface slope (tan θ) and the slope variance s^2 , as shown in equation (3.53):

$$\int p(\theta) d^2 \omega = 1 \Rightarrow p(\theta) = \frac{1}{2\pi s^2 \cos^3 \theta} \exp\left(\frac{-\tan^2 \theta}{2s^2}\right)$$
(3.53)

When used in the scalar microfacet model this leads to the following form for the function f:

$$f = \frac{1}{8\pi s^2 \cos^4 \theta \cos \theta_i \cos \theta_r} \exp\left(\frac{-\tan^2 \theta}{2s^2}\right) R(\Omega)$$
(3.54)

where $R(\Omega)$ is a Fresnel reflectivity function for the selected scalar channel. Note that this form of the BRDF satisfies both the normalization constraint in equation (3.46) and the reciprocity relation in equation (3.49). Note, importantly, that this is a deterministic model. The parameters of the right-hand side are all fixed for a given geometry and surface roughness. The geometry and roughness effects then act as a multiplier to the scattering coefficient—itself a deterministic quantity. Importantly this enables us to extend this model to a fully polarimetric version by replacing $R(\beta)$ by a vector formulation, as follows.

The main complication in formulating a vector BRDF is in matching the local coordinates. There are four such systems to consider. The first is the system defined by the incident wave direction \underline{r}_i and the mean surface normal in the z direction. The second is then defined by the *local* surface normal \underline{n} and incident wave direction. This second is rotated about the incident direction by an angle η_I , as shown in equation (3.55). The third coordinate system is defined by the local surface normal \underline{n} and the reflected wave direction. This has finally to be related to the fourth: namely, that formed by the reflected wave direction and the mean normal z. This is related to the third by a rotation about the reflected direction by an angle η_r , as shown in equation (3.55).

The scattering matrix can then be written as a cascade of three processes: rotation into the local system by η_I , followed by Fresnel scattering, followed by negative rotation out of the local into the global system by η_r . Mathematically this cascade of matrices is shown in equation (3.55), which is a form we have encountered before, being a singular value transformation of the diagonal Fresnel scattering matrix by different left and right singular vectors.

$$\begin{cases} \cos \eta_{i} = \frac{\cos \theta - \cos \Omega \cos \theta_{i}}{\sin \theta_{i} \sin \beta} \\ \cos \eta_{r} = \frac{\cos \theta - \cos \Omega \cos \theta_{r}}{\sin \theta_{r} \sin \Omega} \end{cases} \begin{cases} R_{hh}(\Omega, \varepsilon_{r}) = \frac{\cos \Omega - \sqrt{\varepsilon_{r} - \sin^{2} \Omega}}{\cos \Omega + \sqrt{\varepsilon_{r} - \sin^{2} \Omega}} \\ R_{vv}(\Omega, \varepsilon_{r}) = \frac{\varepsilon_{r} \cos \Omega - \sqrt{\varepsilon_{r} - \sin^{2} \Omega}}{\varepsilon_{r} \cos \Omega - \sqrt{\varepsilon_{r} - \sin^{2} \Omega}} \end{cases} \end{cases}$$
$$\Rightarrow [S] = \sqrt{a_{p}} \begin{bmatrix} \cos \eta_{r} & \sin \eta_{r} \\ -\sin \eta_{r} & \cos \eta_{r} \end{bmatrix}} \cdot \begin{bmatrix} R_{ss}(\Omega) & 0 \\ 0 & R_{pp}(\Omega) \end{bmatrix} \cdot \begin{bmatrix} \cos \eta_{i} & -\sin \eta_{i} \\ \sin \eta_{i} & \cos \eta_{i} \end{bmatrix}} \\ a_{p} = \frac{1}{8\pi s^{2} \cos^{4} \theta \cos \theta_{i} \cos \theta_{r}}} \exp\left(\frac{-\tan^{2} \theta}{2s^{2}}\right)$$
(3.55)

Having determined the form of the [S] matrix, we can now vectorize to obtain the corresponding scattering vector \underline{k} , which can be expressed compactly as a unitary transformation, as shown in equation (3.56):

$$\underline{k}_{M} = \sqrt{a_{p}} \begin{bmatrix} \cos(\eta_{i} - \eta_{r}) & 0 & 0 & \sin(\eta_{i} - \eta_{r}) \\ 0 & \cos(\eta_{i} + \eta_{r}) & \sin(\eta_{i} + \eta_{r}) & 0 \\ 0 & -\sin(\eta_{i} + \eta_{r}) & \cos(\eta_{i} + \eta_{r}) & 0 \\ -\sin(\eta_{i} - \eta_{r}) & 0 & 0 & \cos(\eta_{i} - \eta_{r}) \end{bmatrix} \\ \times \begin{bmatrix} (R_{ss}(\Omega) + R_{pp}(\Omega)) \\ (R_{ss}(\Omega) - R_{pp}(\Omega)) \\ 0 \end{bmatrix}$$
(3.56)

The most important consequence of this model is that all the surface roughness effects are contained in the scalar multiplier a_p . Hence, like the Bragg model for backscatter, there is zero depolarisation under this model—one of the most commonly used for bistatic surface scattering. Consequently the scattering entropy is zero and the scattering coherency matrix is rank 1, as shown in equation (3.57):

$$[T] = \underline{k}_M . \underline{k}_M^{*T} \tag{3.57}$$

There is again a clear discrepancy between the model predictions and observed depolarisation behaviour of real surfaces (see Chapter 9). In particular, as we found with the backscatter Bragg model, there is no way to parameterize smooth change from a zero entropy polariser to Lambertian depolariser. To overcome this a heuristic compromise is often sought by adding a Lambertian component to the microfacet model with parameter a_d , often determined by fitting the model to experimental data. In this case, depolarisation under bistatic surface

scattering can be modelled as shown in equation (3.58):

$$\langle [T] \rangle = \underline{k}_{M} \cdot \underline{k}_{M}^{*T} + a_{d} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$
(3.58)

We investigate the validity of this approach in Chapter 9, when we consider the application of these models to real surface scattering data. Now, however, we turn to consider the second important class of depolarisation problems: namely, those caused by volume scattering.

3.3 Introduction to volume scattering

In addition to wave reflection from boundaries between materials (surface scattering), there is an important second class of polarimetric scattering phenomena to be considered: namely, volume scattering. This occurs within inhomogeneous bulk material that contains local variations in dielectric properties. These act as 'sites' for wave scattering that are distributed in space and act both to increase the scattering cross section and influence propagation of the wave through the medium via the process of wave extinction.

As well as the presence of material inhomogeneities, volume scattering also requires significant penetration of the wave into the medium to excite these secondary scattering centres. This latter observation leads us to a simple test to assess the relative importance of volume versus surface scattering, based on the effective complex permittivity of a medium ε_r . As shown in equation (3.59), and as a generalization of the concept of skin depth from equation (3.13), we can adopt an exponential structure function assumption in the material, and use this to define a significant power penetration depth δ_p as shown:

$$\frac{\varepsilon_r = \varepsilon' - i\varepsilon''}{n = \sqrt{\varepsilon_r} = n' - in''} \Rightarrow e^{i\omega(t - \frac{n\varepsilon}{c})} = e^{i\omega t} e^{-i\frac{\omega n}{c}z} = e^{i\omega t} e^{-i\frac{\omega n'}{c}z} e^{-\frac{\omega n''}{c}z}$$

$$\frac{P(z = \delta_p)}{P(z = 0)} = \frac{1}{e} \Rightarrow \delta_p = \frac{\lambda_o}{2\pi n''} = \frac{\lambda_o}{2\pi \left|\operatorname{Im}(\sqrt{\varepsilon_r})\right|}$$
(3.59)

We see that the penetration can be expressed as a scale factor times the free space wavelength, and for volume scattering to be possible we require this factor to be much greater than unity. We see that the factor depends only on the imaginary part of the square root of the effective dielectric constant, but often a further simplification is made, since for most materials in microwave remote sensing the imaginary part is much smaller than the real. We therefore have the following common approximation:

$$\frac{\varepsilon''}{\varepsilon'} < 0.1 \to \delta_p \approx \frac{\lambda_o}{2\pi} \frac{\sqrt{\varepsilon'}}{\varepsilon''} \tag{3.60}$$

To illustrate this relation, we now consider five important application areas in microwave remote sensing: namely, scattering from water surfaces, ice, snow, soil and vegetation (Ulaby, 1982, 1986; Dobson, 1985).

Water

Pure water has a high dielectric constant (in both real and imaginary parts; see equation (3.15)) across the microwave spectrum, and this severely limits the penetration depth of microwaves. Radar scattering by water can therefore be considered a surface-dominated problem, albeit often a dynamic one with time varying and spatially inhomogeneous roughness due to interacting wind and current systems. Note, however, that breaking waves pose an important example where volume scattering can occur in the water context, as these generate a turbulent air/water mixture with a relatively low effective dielectric constant and where wave penetration can be significant and volume scattering effects be observed.

Land and sea ice

Pure ice has a real dielectric constant around $\varepsilon_r = 3.15$, and a low loss-tangent, so penetration depths can extend 300 wavelengths or more, depending on frequency. In this case, volume scattering can be significant in thick ice deposits such as occur in land ice. Sea ice generally has a much larger imaginary part due to the presence of brine, with first-year ice ($\varepsilon_r \approx 3.3 - i0.25$) having a higher loss than brine-deficient multi-year ice ($\varepsilon_r \approx 3.3 - i0.03$). Penetration depths are therefore higher into multi-year ice, and vary from 10 down to 1 wavelength.

Snow

Dry snow is a mixture of ice and air—both very low-loss materials—and so penetration can occur to hundreds of wavelengths, with a small overall dielectric constant lying in the range 1.5–2. The imaginary part of the dielectric constant is, however, very sensitive to changes in water content, and so wet snow (which has dielectric constant always less than 4) shows a greater variety of penetration depths. In general, therefore, snow can behave as a surface or volume scatterer (or both), depending on environmental conditions, viewing geometry, and operating frequency.

Soils

As discussed in equation (3.14), soil is a composite material, the dielectric constant of which depends on a large number of parameters. It is therefore difficult to derive general conclusions about soil penetration. However, across a large part of the microwave spectrum, penetration depth depends primarily on soil moisture content, and for most mid-latitude soil types penetration is therefore seldom more than a wavelength or so. An important exception must be made for hyper arid sand/soil environments, where penetrations of many wavelengths can occur. Soil is therefore similar to snow in that both surface and volume scattering can occur, although in most mid-latitude agricultural areas the assumption is often made of pure surface scattering.

Vegetation

Vegetation scattering is perhaps the most complex of all, and simple models for penetration depth are often of limited use. Indeed, vegetation scattering provides an important focus for the development of polarimetric interferometry in Chapter 6, which allows us, amongst other things, to estimate the penetration depth into vegetation. It is sufficient here to point out that vegetation is often a strong volume scattering environment, combining multiple wavelengths of penetration with the presence of strong dielectric discontinuities (moist branches, twigs, leaves, and so on), which are often larger than a wavelength, and act to scatter the waves in a complex manner.

Surface and volume scattering effects are often discriminated on the basis of their dependence on angle of incidence. Surface scattering generally shows a fall-off with increasing angle, while volume effects vary only slightly with change of angle. However, in this text we are concerned primarily with the polarisation properties of the two phenomena. In preparation for a discussion of the polarisation properties of volume effects, we now consider scattering by an individual particle. In the next section we use these results with averaging to analyse the depolarisation behaviour of random volume scattering.

3.3.1 Small particle scattering

In the previous sections we dealt with reflection and scattering at a rough interface between two media. While this provides a good model for surfaces, it does not deal with the case of volume scattering. This occurs when we have a three-dimensional distribution of scattering centres, and arises in practice for a description of vegetation cover and in soil and ice penetration. One approach to modelling such effects is to consider each localized scattering centre as a particle of known shape and material composition embedded in an homogeneous background. The problem then reduces to establishing the scattered fields by the particle, and combining a large number of such elements to provide a macroscopic description of the scattering medium.

In general, even the first stage of this process can be complicated, as Maxwell's equations must be solved across the boundary of the scatterer. For spherical particles this can be achieved using an expansion of incident and scattered fields in spherical wave functions, and matching of the fields in a similar manner to that used in the Fresnel equations. This leads to a convergent solution known as the Mie series, which can be used for calculating the full scattering matrix of spheres of arbitrary size and material composition and for arbitrary scattering angle Ψ (van de Hulst, 1981; Hovenier, 2004). However, spherical particles have a very special symmetry and hence limited polarisation response (although they can still depolarise via multiple scattering, as demonstrated in Chapter 9). Furthermore, they are not found in many natural media such as vegetation, and consequently such a symmetric solution is of limited applicability in radar remote sensing.

A more useful approximation is to assume spheroidal particles, with one long axis and two equal minor axes, as shown in Figure 3.20 (Jin, 1994a). Here we



Fig. 3.20 Spheroidal particle geometry and particle coordinate frame

can now model variations in shape from prolate (needle-like) through spherical and into oblate (disc-like). This provides a more realistic variation of shape for modelling in many natural media. Again the solution can be established from Maxwell's equations by expanding fields in terms of vector spheroidal wave functions and matching coefficients across the boundary. However, this solution is not nearly as convenient as the Mie series, for the following reasons (Mishchenko, 2000, 2006, 2007):

- The computation of vector spheroidal functions is a complicated mathematical and numerical problem in itself, especially for absorbing particles.
- The spheroidal functions are not orthogonal on the surface of the scatterer, and the unknown expansion coefficients must therefore be solved using an infinite set of algebraic equations, which in practice must be truncated and solved numerically. Unfortunately, for large particles this system becomes ill-conditioned, and care is needed in the numerical computation of results.

As a result of these difficulties, several alternative numerical methods have been developed based, for example, on finite element, finite difference, and Tmatrix approximations. These provide more robust general-purpose numerical solutions, but are too complicated in structure for simple analytical studies.

To avoid these problems we illustrate the important polarisation physics by considering the case of Rayleigh scattering by spheroids. Lord Rayleigh (1842–1919) first derived an approximation for scattering in the small particle limit when $\beta a \ll 1$ by assuming that the fields in and near the particle are the same as those derived from electrostatics, and that the internal field inside the particle is homogeneous. This approximation leads to a fully analytic solution, even for chiral particles, which as we have seen in Chapter 1, case III, is a case of special importance in polarimetric studies. We now consider the main polarimetric features of this solution method.

With reference to Figure 3.20, we now assume that the particle dimensions are small compared to the wavelength of the incident radiation so that the field inside may be considered constant, and we can ignore any time delays for fields to propagate across the particle. In quantitative terms this requires the following condition:

$$|n|\,\beta a \ll 1 \tag{3.61}$$

where *n* is the complex refractive index of the particle, β is the wavenumber $2\pi/\lambda$, and *a* is the volume equivalent sphere radius of the particle. The first important stage in the solution is to recognize that the particle has a definite orientation, specified by its major axis, shown as the dotted line in Figure 3.20. This can be used to define a particle frame or coordinate system that has its z axis aligned with the major axis.

The incident field induces a dipole moment in the particle, which then acts to reradiate or scatter energy. We wish to allow the possibility of chiral particles (such as a small helix) when the incident electric field can generate circulating as well as linear currents. This means that we must allow both electric (\underline{p}) and magnetic (\underline{m}) dipole moments in the particle (see equations (1.10) and (1.11)). These moments will evidently depend on the orientation of the driving

field, and thus are to be represented by polarizability tensors. The moment vectors are then related to the driving field vectors by the following matrix equation:

$$\begin{bmatrix} \underline{p} \\ \underline{\underline{m}} \end{bmatrix} = \begin{bmatrix} \overline{\overline{\alpha}}_{ee} & \overline{\overline{\alpha}}_{em} \\ -\overline{\overline{\alpha}}_{em} & \overline{\overline{\alpha}}_{mm} \end{bmatrix} \cdot \begin{bmatrix} \underline{\underline{E}} \\ \underline{\underline{H}} \end{bmatrix}$$
(3.62)

where $\overline{\alpha}_{ij}$ are 3 × 3 polarizability tensors, which are all diagonal in the particle frame. The field radiated by elementary electric (*p*) and magnetic (*m*) dipole moments can be obtained directly from Maxwell's equations, as shown in equations (1.10) and (1.11), and are given explicitly in equation (3.63):

$$\underline{\underline{E}} = \left[\omega^2 \mu_o (\overline{\overline{\underline{I}}} - \underline{rr}) \underline{\underline{p}} - \omega \beta_o \underline{\underline{r}} \times \underline{\underline{m}}\right] \frac{\exp\left(-i\beta_o r\right)}{4\pi r}$$
(3.63)

Here <u>r</u> is a unit vector in the direction of the scattered wave. Before we can use this relationship with equation (3.62) we first must transform the coordinate system from the particle into the laboratory frame. The latter can be defined as shown in Figure 3.21. Without loss of generality, we consider the wave incident along the +z direction and scattering in the y-z plane in the direction <u>r</u>, which makes an angle Ψ with the z axis. We then wish to find the 2 × 2 scattering amplitude matrix [S] for radiation polarised parallel (P) or perpendicular (S) to the scattering plane, as shown in Figure 3.21. Using the <u>r</u> vector as defined in Figure 3.21 together with equation (3.63), we can express the radiated field by the particle as shown in equation (3.64):

$$E_{S'} = \left[\frac{1}{\varepsilon_o}p_1 - \frac{1}{\sqrt{\varepsilon_o\mu_o}}\left(m_3\sin\psi - m_2\cos\psi\right)\right] \frac{\beta_o^2\exp(-i\beta_o r)}{4\pi r}$$

$$E_{P'} = \left[\frac{1}{\varepsilon_o}\left(p_2\cos\psi - p_3\sin\psi\right) - \frac{1}{\sqrt{\varepsilon_o\mu_o}}m_1\right] \frac{\beta_o^2\exp(-i\beta_o r)}{4\pi r}$$
(3.64)



Fig. 3.21 Laboratory coordinate frame for particle scattering

To solve this we need the components of $\underline{p}^{\text{lab}} = (p_1, p_2, p_3)$ and $\underline{m}^{\text{lab}} = (m_1, m_2, m_3)$. To convert the polarizability tensors from the particle to the laboratory frame we make use of Euler angles relating one coordinate system to the other (Goldstein, 1980). As we are dealing with spheroids, the transformation from one to the other requires only two angles, and we choose a particle orientation angle θ and a tilt angle τ as shown in equation (3.65):

$$[R] = \begin{bmatrix} \cos \tau & 0 & -\sin \tau \\ 0 & 1 & 0 \\ \sin \tau & 0 & \cos \tau \end{bmatrix} \cdot \begin{bmatrix} \cos \theta & \sin \theta & 0 \\ -\sin \theta & \cos \theta & 0 \\ 0 & 0 & 1 \end{bmatrix}$$
$$= \begin{bmatrix} \cos \tau \cos \theta & \cos \tau \sin \theta & -\sin \tau \\ -\sin \theta & \cos \theta & 0 \\ \sin \tau \cos \theta & \sin \tau \sin \theta & \cos \tau \end{bmatrix}$$
(3.65)

Using this rotation matrix we can then express equation (3.62) in the laboratory frame, as shown in equation (3.66):

$$\begin{bmatrix} \underline{p}^{lab} \\ \underline{\underline{m}}^{lab} \end{bmatrix} = \begin{bmatrix} R^{-1}\overline{\overline{\alpha}}_{ee}R & R^{-1}\overline{\overline{\alpha}}_{em}R \\ -R^{-1}\overline{\overline{\alpha}}_{em}R & R^{-1}\overline{\overline{\alpha}}_{mm}R \end{bmatrix} \cdot \begin{bmatrix} \underline{\underline{E}}^{inc} \\ \underline{\underline{H}}^{inc} \end{bmatrix}$$
(3.66)

where now we can make use of the incident electric and magnetic fields for a TEM wave in the laboratory frame directly as

$$\underline{E}^{inc} = \begin{bmatrix} E_S & E_P & 0 \end{bmatrix}^T \Rightarrow \underline{H}^{inc} = \sqrt{\frac{\varepsilon_o}{\mu_o}} \begin{bmatrix} -E_P & E_S & 0 \end{bmatrix}$$
(3.67)

Evaluation of the elements of one block of equation (3.66) will illustrate the general form of the solution, as shown in equation (3.68):

$$R^{-1}\overline{\alpha}R = \begin{bmatrix} \cos\tau\cos\theta & -\sin\theta & \sin\tau\cos\theta\\ \cos\tau\sin\theta & \cos\theta & \sin\tau\sin\theta\\ -\sin\tau & 0 & \cos\tau \end{bmatrix} \cdot \begin{bmatrix} \rho_1 & 0 & 0\\ 0 & \rho_2 & 0\\ 0 & 0 & \rho_2 \end{bmatrix} \cdot \begin{bmatrix} \cos\tau\cos\theta & \cos\tau\sin\theta & -\sin\tau\\ -\sin\theta & \cos\theta & 0\\ \sin\tau\cos\theta & \sin\tau\sin\theta & \cos\tau \end{bmatrix}$$
$$= \begin{bmatrix} \rho_2 + (\rho_1 - \rho_2)\cos^2\tau\cos^2\theta & (\rho_1 - \rho_2)\cos^2\tau\sin\theta\cos\theta & -(\rho_1 - \rho_2)\cos\tau\sin\tau\cos\theta\\ (\rho_1 - \rho_2)\cos^2\tau\sin\theta\cos\theta & \rho_2 + (\rho_1 - \rho_2)\cos^2\tau\sin^2\theta & -(\rho_1 - \rho_2)\cos\tau\sin\tau\sin\theta\\ - (\rho_1 - \rho_2)\cos\tau\sin\tau\cos\theta & -(\rho_1 - \rho_2)\cos\tau\sin\tau\sin\theta & \rho_1\sin^2\tau + \rho_2\cos^2\tau \end{bmatrix}$$
(3.68)

We now have all the elements we require for a solution. We start by setting the incident field as S or P polarised in equation (3.67), and then use equation (3.66) with (3.68) for each block element to obtain the <u>p</u> and <u>m</u> vectors. Finally we use equation (3.64) to derive the S and P components of the scattered field. In this way we can calculate the scattering matrix as a function of the three angles Ψ , θ , τ and the four pairs of polarizability constants for the particle.

Rather than calculate the full set of equations, we consider instead two important special cases that permit further simplification of the above equations.

3.3.1.1 Example 1: scattering by small non-chiral particles

In this case we can set $\overline{\overline{\alpha}}_{em} = \overline{\overline{\alpha}}_{me} = \overline{\overline{\alpha}}_{mm} = \overline{\overline{0}}$, and the equations simplify to the following form:

$$p_{1} \\ p_{2} \\ p_{3} \end{bmatrix} = \begin{bmatrix} \rho_{2} + (\rho_{1} - \rho_{2})\cos^{2}\tau\cos^{2}\theta & (\rho_{1} - \rho_{2})\cos^{2}\tau\sin\theta\cos\theta & -(\rho_{1} - \rho_{2})\cos\tau\sin\tau\cos\theta \\ (\rho_{1} - \rho_{2})\cos^{2}\tau\sin\theta\cos\theta & \rho_{2} + (\rho_{1} - \rho_{2})\cos^{2}\tau\sin^{2}\theta & -(\rho_{1} - \rho_{2})\cos\tau\sin\tau\sin\theta \\ -(\rho_{1} - \rho_{2})\cos\tau\sin\tau\cos\theta & -(\rho_{1} - \rho_{2})\cos\tau\sin\tau\sin\theta & \rho_{1}\sin^{2}\tau + \rho_{2}\cos^{2}\tau \end{bmatrix} \cdot \begin{bmatrix} E_{p}^{inc} \\ E_{p}^{inc} \\ -E_{p}^{inc} \\ 0 \end{bmatrix}$$

$$\Rightarrow E_{S'} = \begin{bmatrix} \frac{1}{\varepsilon_{o}}p_{1} \end{bmatrix} \frac{\beta_{o}^{2}\exp(-i\beta_{o}r)}{4\pi r}$$

$$E_{p'} = \begin{bmatrix} \frac{1}{\varepsilon_{o}}(p_{2}\cos\psi - p_{3}\sin\psi) \end{bmatrix} \frac{\beta_{o}^{2}\exp(-i\beta_{o}r)}{4\pi r}$$

$$(3.69)$$

From these equations the elements of the scattering matrix can then be directly obtained as shown in equation (3.70):

$$[S] = \begin{bmatrix} S_{PP} & S_{SP} \\ S_{PS} & S_{SS} \end{bmatrix}$$
$$\Rightarrow \begin{cases} S_{PP} = (A_P - 1)\sin\theta\cos\tau(\sin\theta\cos\tau\cos\Psi - \sin\tau\sin\Psi) + \cos\Psi \\ S_{SP} = (A_P - 1)\cos\theta\cos\tau(\sin\theta\cos\tau\cos\Psi - \sin\tau\sin\Psi) \\ S_{PS} = (A_P - 1)\sin\theta\cos\theta\cos^{2}\tau \\ S_{SS} = (A_P - 1)\cos^{2}\theta\cos^{2}\tau + 1 \end{cases}$$
(3.70)

where we have defined the particle anisotropy A_p as a ratio of principal polarizabilities, so that

$$A_p = \frac{\rho_1}{\rho_2} \tag{3.71}$$

For simplicity we have cancelled all factors common to the elements of [S]. Of particular interest is the case of backscatter ($\Psi = 180^{\circ}$), when the equations further simplify to those shown in equation (3.72):

$$[S] = \begin{bmatrix} S_{PP} & S_{SP} \\ S_{PS} & S_{SS} \end{bmatrix} \Rightarrow \begin{cases} S_{PP} = -(A_P - 1)\sin^2\theta\cos^2\tau - 1 \\ S_{SP} = -(A_P - 1)\sin\theta\cos\theta\cos^2\tau \\ S_{PS} = (A_P - 1)\sin\theta\cos\theta\cos^2\tau \\ S_{SS} = (A_P - 1)\cos^2\theta\cos^2\tau + 1 \end{cases}$$
(3.72)

This scattering matrix is expressed in the wave coordinates, and if we convert to the sensor (BSA) system, commonly used in radar studies, we obtain the following solution:

$$[S]_{BSA} = \begin{bmatrix} S_{PP} & S_{SP} \\ S_{PS} & S_{SS} \end{bmatrix} \Rightarrow \begin{cases} S_{PP} = (A_P - 1)\sin^2\theta\cos^2\tau + 1\\ S_{SP} = (A_P - 1)\sin\theta\cos\theta\cos^2\tau\\ S_{PS} = (A_P - 1)\sin\theta\cos\theta\cos^2\tau\\ S_{SS} = (A_P - 1)\cos^2\theta\cos^2\tau + 1 \end{cases}$$
(3.73)

which we see is complex symmetric, as expected. Note that this matrix has a straightforward eigenvalue expansion, as shown in equation (3.74):

$$[S]_{BSA} = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 \\ 0 & \sin^2\tau + A_P\cos^2\tau \end{bmatrix} \cdot \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$$
(3.74)

which demonstrates that from measurement of the backscatter amplitude matrix of a single particle we can determine the angle θ from the singular vectors of [S], and from the ratio of singular values we obtain $\kappa = A_P \cos^2 \tau + \sin^2 \tau$. This ratio is 1 when $\tau = 90^\circ$; that is, when viewed along its major axis, in which case it has a circular symmetric cross-section. In this case the particle is indistinguishable from a sphere. However, when $\tau = 0$ —when the major axis lies in the plane of polarisation—we obtain a ratio equal to A_P , the particle anisotropy.

Interpretation of A_P in terms of the particle geometry requires evaluation of the ratio of principal values of polarizability. This can be evaluated analytically for small spheroids of volume V (Ishimaru, 1991; Tsang, 1985), and yields the following result:

$$\rho_i = \frac{\varepsilon_o \left(\varepsilon_r - 1\right) V}{1 + \left(\varepsilon_r - 1\right) L_i} \Rightarrow A_P = \frac{L_2 + \frac{1}{\varepsilon_r - 1}}{L_1 + \frac{1}{\varepsilon_r - 1}}$$
(3.75)

where L_1 and L_2 are shape functions given explicitly by equation (3.76):

$$L_{1} = \begin{cases} \frac{1-e^{2}}{e^{2}} \left(-1+\frac{1}{2e} \ln \frac{1+e}{1-e}\right) & x_{1} > x_{2} = x_{3} \quad e^{2} = 1-\frac{x_{2}^{2}}{x_{1}^{2}} \quad (prolate) \\ \frac{1+f^{2}}{f^{2}} \left(1-\frac{1}{f} \arctan f\right) & x_{1} < x_{2} = x_{3} \quad f^{2} = \frac{x_{2}^{2}}{x_{1}^{2}} - 1 \quad (oblate) \\ L_{2} = L_{3} = \frac{1}{2}(1-L_{1}) \quad (3.76) \end{cases}$$

Figure 3.22 shows a plot of the relationship between particle shape $r = x_1/x_2$ and the polarizability ratio for low and high values of dielectric constant. Hence we see that A_p is a general indicator of particle shape, although it is bounded in value by the dielectric constant to lie between limits given by equation (3.77)



Fig. 3.22 Relationship between polarizability ratio and shape for varying dielectric constant

(Ablitt, 2000):

$$\frac{1}{\varepsilon_r} < A_P < \frac{\varepsilon_r + 1}{2} \tag{3.77}$$

Note also that to a good approximation $\frac{L_1}{L_2} \approx \frac{x_2}{x_1}$ (the error is less than 0.04 in all cases), and hence for large ε_r equation (3.75) yields $A_P \approx \frac{x_1}{x_2}$, which then becomes a direct indicator of particle shape. This again demonstrates the importance of SVD in the analysis of the scattering matrix. In this case the singular vectors lead to estimation of the particle orientation angle and the ratio of singular values to the projected shape and dielectric constant of the particle.

3.3.1.2 Example 2: scattering by small chiral particles

As a second example of the small particle scattering matrix, we now consider spheroids made from chiral material. In this case the magnetic dipole moment contribution must be included in the radiated field. However, we can still make some useful simplifying assumptions.

We introduced chiral materials in equation (1.59), and then defined the scalar chiral admittance in equation (1.60). We noted that chirality is very weak for many natural materials. Under this assumption we can assume that the parameter $\beta_0 \Omega$ is small, as suggested in equation (1.69). This then has the following implications for the polarizability tensors in equation (3.52):

- The magnetic tensor can be set to zero $\overline{\overline{\alpha}}_{mm} = \overline{\overline{0}}$, as the magnetic field effects will be small compared to those induced by the electric field.
- We can ignore any chirality effects in the electric field tensor $\overline{\overline{\alpha}}_{ee}$, and hence this tensor has the same form as in example 1 for non-chiral particles.
- We must consider the effect of the mixed tensor $\overline{\overline{\alpha}}_{em}$, which in the simplest case will be linearly dependent on the small parameter $\beta_0 \Omega$.

With these comments in mind we postulate the following form of equation (3.62):

$$\begin{bmatrix} \underline{p}^{lab} \\ \underline{\underline{m}}^{lab} \end{bmatrix} = \begin{bmatrix} R^{-1} \overline{\overline{\alpha}}_{ee} R & R^{-1} \overline{\overline{\alpha}}_{em} R \\ -R^{-1} \overline{\overline{\alpha}}_{em} R & \overline{\overline{0}} \end{bmatrix} \cdot \begin{bmatrix} \underline{\underline{E}}^{inc} \\ \underline{\underline{H}}^{inc} \end{bmatrix} \Rightarrow \begin{cases} \overline{\overline{\alpha}}_{ee} = \begin{bmatrix} \rho_1 & 0 & 0 \\ 0 & \rho_2 & 0 \\ 0 & 0 & \rho_2 \end{bmatrix} \\ \overline{\overline{\alpha}}_{em} = \begin{bmatrix} \delta_1 & 0 & 0 \\ 0 & \delta_2 & 0 \\ 0 & 0 & \delta_2 \end{bmatrix}$$
(3.78)

From the assumption of a linear dependence on $\beta_0 \Omega$ we also have

$$A_P = \frac{\rho_1}{\rho_2} = \frac{\delta_1}{\delta_2} \quad i\kappa = \frac{\delta_1}{\alpha_1} = i\frac{\sqrt{\varepsilon_r}}{\varepsilon_r - 1}\beta_0\Omega \tag{3.79}$$

where κ is a dimensionless chirality parameter. Note that the $\overline{\overline{\alpha}}_{em}$ tensor is purely imaginary, as pointed out in equation (1.60). This allows us to rewrite

equation (3.78) in the simplified form shown in equation (3.80):

$$\begin{bmatrix} \underline{p}^{lab} \\ \underline{\bar{m}}^{lab} \end{bmatrix} = \begin{bmatrix} R^{-1} \overline{\overline{\alpha}}_{ee} R & R^{-1} \overline{\overline{\alpha}}_{em} R \\ -R^{-1} \overline{\overline{\alpha}}_{me} R & \overline{\overline{0}} \end{bmatrix} \cdot \begin{bmatrix} \underline{\underline{E}}^{inc} \\ \underline{\underline{H}}^{inc} \end{bmatrix} \Rightarrow \begin{cases} \overline{\overline{\alpha}}_{ee} = \begin{bmatrix} A_P & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \\ \overline{\overline{\alpha}}_{em} = i\kappa \begin{bmatrix} A_P & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$
(3.80)

We can now evaluate this explicitly. The result is shown in equation (3.81):

$$\begin{split} [S] &= \begin{bmatrix} S_{PP} & S_{SP} \\ S_{PS} & S_{SS} \end{bmatrix} \\ \\ \Rightarrow \begin{cases} S_{PP} &= (A_P - 1)\sin\theta\cos\tau\,(\sin\theta\cos\tau\cos\psi + \sin\tau\sin\Psi) + \cos\Psi \\ &+i\kappa\,(A_P - 1)\cos\theta\cos\tau\,(\sin\theta\cos\tau - \cos\tau\sin\theta\cos\Psi - \sin\tau\sin\Psi) \\ S_{SP} &= (A_P - 1)\cos\theta\cos\tau\,(\sin\theta\cos\tau\cos\Psi + \sin\tau\sin\Psi) \\ &+i\kappa[(A_P - 1)\cos\tau\,(\cos^2\theta\cos\tau + \cos\tau\sin^2\theta\cos\Psi + \sin\tau\sin\theta\sin\Psi) + (1+\cos\Psi)] \\ S_{PS} &= (A_P - 1)\sin\theta\cos\theta\cos^2\tau \\ &-i\kappa[(A_P - 1)\cos\tau\,(\cos^2\theta\cos\tau + \cos\tau\sin^2\theta\cos\Psi - \sin\tau\sin\theta\sin\Psi) + (1+\cos\Psi)] \\ S_{SS} &= (A_P - 1)\cos^2\theta\cos^2\tau + 1 \\ &+i\kappa\,(A_P - 1)\cos\theta\cos\tau\,(\sin\theta\cos\tau - \cos\tau\sin\theta\cos\Psi - \sin\tau\sin\Psi) \\ \end{cases}$$
(3.81)

Again, as we are primarily interested in backscatter, we set $\Psi = 180^{\circ}$. At the same time we apply the sensor coordinate corrections to obtain the backscatter matrix in the BSA convention, as shown in equation (3.82) (Cloude, 2002b):

$$[S]_{BSA} = \begin{bmatrix} S_{PP} & S_{SP} \\ S_{PS} & S_{SS} \end{bmatrix} \Rightarrow \begin{cases} S_{PP} = (A_P - 1)\sin^2\theta\cos^2\tau + 1 \\ -i\kappa (A_P - 1)\sin 2\theta\cos^2\tau \\ S_{SP} = (A_P - 1)\sin\theta\cos\theta\cos^2\tau \\ -i\kappa (A_P - 1)\cos^2\tau\cos 2\theta \\ S_{PS} = (A_P - 1)\sin\theta\cos\theta\cos^2\tau \\ -i\kappa (A_P - 1)\cos^2\tau\cos 2\theta \\ S_{SS} = (A_P - 1)\cos^2\theta\cos^2\tau + 1 \\ +i\kappa (A_P - 1)\sin 2\theta\cos^2\tau \end{cases}$$
(3.82)

This can be written as a decomposition of the scattering matrix in the following form:

$$[S] = [S]_A + [S]_B = R(-\theta) (S_1 + S_2) R(\theta)$$

= $R(-\theta) \cdot \left(\begin{bmatrix} 1 & 0 \\ 0 & A_P \cos^2 \tau + \sin^2 \tau \end{bmatrix} - \kappa (A_P - 1) \cos^2 \tau \begin{bmatrix} 0 & i \\ i & 0 \end{bmatrix} \right) \cdot R(\theta)$
(3.83)

This result shows that the presence of a small amount of chirality in the particle leads to a decomposition of the backscattering matrix into the sum of two terms: S_A , the matrix for non-chiral particles; and S_B , a chiral perturbation which is manifest as a scalar multiple of one of the Pauli spin matrices.

3.3.2 Scattering by large particles

As particle size increases compared to the wavelength, then the scattering behaviour changes, with forward scattering dominating over backscatter. In this section we consider some of the implications for the polarisation properties of large particle scattering. Detailed analytical calculations are generally difficult for large particles, and resort must usually be made to numerical and approximation techniques such as the T-matrix and discrete dipole approximation (DDA) (Mishchenko, 2000, 2006, 2007). Public domain codes are available for both these advanced techniques; for DDA see http://www.science.uva.nl/research/scs/Software/adda, and for the T-matrix see http://www.giss.nasa.gov/~crmim/t_matrix.html. These codes can cope not only with single particles, but also with small aggregates such as multi-sphere clusters, accounting for all multiple interactions between particles.

One problem that does allow some analytic insight is that for scattering by spheres of arbitrary size and material. Small spheres of radius $r << \lambda$ act as Rayleigh scatterers, characterized by a polarizability ρ , as considered in the previous section. They have a bistatic amplitude [S] matrix with the characteristic dipole radiation pattern, a function of the scattering angle Ψ (where $\Psi = 0^{\circ}$ is forward scatter, and $\Psi = 180^{\circ}$ is backscatter), as shown in equation (3.84):

$$[S] = i\beta^3 \rho \begin{bmatrix} \cos \Psi & 0\\ 0 & 1 \end{bmatrix}$$
(3.84)

Here VV (where V is perpendicular to the scattering plane) has a uniform scattering pattern, showing equal back and forward scatter. HH, on the other hand, shows a null for $\Psi = 90^{\circ}$ —for lateral scattering—but maintains equality of forward and backscattering amplitudes. As particle radius increases, two things happen:

- 1. Because of the spherical symmetry, the crosspolarisation terms of [S] remain zero; that is, the matrix is always diagonal in the H and V basis (parallel and perpendicular to the scattering plane). Therefore, scattering by a sphere has SU(2) transformation properties (any complex combination of HH and VV can be formed by two-element unitary weight vectors \underline{w}). Therefore, scattering by a sphere of arbitrary size and material, scattering into an arbitrary direction, can be mapped as a point on the scattering sphere of Figure 2.8.
- 2. The HH and VV scattering amplitudes become functions of Ψ, and are complex (they have a phase difference and amplitude ratio different from unity). For forward and backscatter cases, however, symmetry forces the ratio to be ±1, with sign depending on coordinates used. Also, in general terms the ratio of forward scattering to backscattering increases with particle size, and there appears a more complicated pattern of nulls in the scattering diagram.

In 1908, Gustav Mie (1869–1957), in a milestone paper, used Maxwell's equations, together with the appropriate boundary conditions, to solve for these two complex functions for arbitrary sphere size and material. The solution is expressed as an infinite series, with so-called Mie coefficients a_n and b_n , as shown in equation (3.85) (Hovenier, 2004; Ishimaru, 1991):

$$[S] = \begin{bmatrix} S_{HH}(\Psi) & 0\\ 0 & S_{VV}(\Psi) \end{bmatrix}$$

$$\Rightarrow \begin{cases} S_{VV}(\Psi) = \sum_{n=1}^{\infty} \frac{2n+1}{n(n+1)} \{a_n \pi_n (\cos \Psi) + b_n \tau_n (\cos \Psi)\} \\ S_{HH}(\Psi) = \sum_{n=1}^{\infty} \frac{2n+1}{n(n+1)} \{b_n \pi_n (\cos \Psi) + a_n \tau_n (\cos \Psi)\} \end{cases}$$
(3.85)

where the component functions π_n and τ_n are defined as follows:

$$\pi_n(\cos\Psi) = \frac{dP_n(\cos\Psi)}{d\cos\Psi}$$

$$\tau_n(\cos\Psi) = \cos\Psi\pi_n(\cos\Psi) - \sin^2\Psi\frac{d\pi_n(\cos\Psi)}{d\cos\Psi}$$
(3.86)

where $P_n(..)$ are the Legendre polynomials (see equation (5.49)). The Mie coefficients a_n and b_n depend on the (generally complex) refractive index and size parameter $x = \beta r$, where *r* is the radius of the sphere. The computation of these coefficients can be quite laborious, especially for large particles, and there are now available several public domain codes for their numerical calculation.

An example dataset (courtesy of Michael Mishchenko of NASA Goddard) for a sphere of size $\beta r = 4$ and refractive index n = 1.32 (equivalent to water/ice at optical wavelengths) is shown in Figures 3.23–3.25. Figure 3.23 shows the phase function—a rather ambiguous term, as it contains no relation to phase between complex numbers. Rather, it is a plot of the 1,1 element of the Mueller matrix [*M*]. From equation (2.48) we see that this is just the total scattered power (trace of coherency matrix). Figure 3.23 therefore shows how the total scattered power varies with scattering angle. We note the enhanced scattering in the



Fig. 3.23 Phase function for scattering by a large dielectric sphere ($\beta r = 4$, n = 1.32)



Fig. 3.24 Amplitude of scattering matrix elements S_{HH} and S_{VV} for scattering by a large dielectric sphere ($\beta r = 4, n = 1.32$)

Fig. 3.25 Scattering sphere representation for a large dielectric sphere ($\beta r = 4$, n = 1.32)

forward direction ($\Psi = 0^{\circ}$) compared to backscatter ($\Psi = 180^{\circ}$). This is due to the large electrical size of the particle. From a polarisation point of view we are more interested in the two complex numbers from the diagonal of the amplitude matrix. For example, Figure 3.24 shows how the dB amplitudes of the two [S] matrix elements vary with scattering angle. Here we note important differences, showing variation of polarised structure with angle. (Note, however, that for forward scatter and backscatter the amplitude ratio is 0 dB, as expected from symmetry.) There is still an additional parameter to consider, however: the phase between HH and VV. The combined amplitude and phase variations can be jointly visualized as points on the scattering sphere, as shown in Figure 3.25. On the left is a three-dimensional representation of the variation of the [S] matrix with scattering angle for the forward scattering hemisphere. On the right is a polar projection of this sphere, which allows a more quantitative interpretation. For example, we can see the scattering matrix start on the equator on the left side for backscattering, and then progress with scattering angle before spiralling towards the antipodal point on the right-hand side for forward scatter.

We have shown by example that even spherical scatterers can generate interesting polarisation behaviour in bistatic scattering. When we extend this to consider non-spherical particles such as spheroids and even more complicated geometries, then the polarisation response changes even more (in particular for backscatter, which is useful for radar studies). However, there is one issue we have yet to consider: that such particles are usually not viewed alone, but in clouds with some distribution over size, shape, orientation, and so on. It is then of interest to see how much of this structure survives the averaging inherent in such depolarisation processes. We now turn to consider such effects in volume scattering.

3.4 Depolarisation in volume scattering

As an important example of residual information retrieval in the presence of depolarisation with azimuthal symmetry, we now consider backscattering by a random cloud of small spheroidal particles. The [S] matrix for an individual such particle was derived in equation (3.74), and here we start by reformulating the [S] matrix as a scattering vector transformation as shown in equation (3.87):

$$\underline{k} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos 2\theta & -\sin 2\theta \\ 0 & \sin 2\theta & \cos 2\theta \end{bmatrix} \begin{bmatrix} 2 + X \cos^2 \tau \\ X \cos^2 \tau \\ 0 \end{bmatrix}$$
$$= \begin{bmatrix} 2 + X \cos^2 \tau \\ X \cos^2 \tau \cos 2\theta \\ X \cos^2 \tau \sin 2\theta \end{bmatrix}$$
(3.87)

where $X = (A_P - 1)$, and A_P is ratio of particle polarizabilities—itself a function of particle shape and dielectric constant. The angles θ and τ dictate the orientation of the spheroid. Now consider a cloud of such particles with random orientation. In this case we must first consider the coherency matrix of an individual particle, as shown in equation (3.88):

$$[T] = \begin{bmatrix} |2 + X\cos^{2}\tau^{2}|^{2} & (2 + X\cos^{2}\tau)X^{*}\cos^{2}\tau\cos 2\theta & (2 + X\cos^{2}\tau)X^{*}\cos^{2}\tau\sin 2\theta \\ (2 + X^{*}\cos^{2}\tau)X\cos^{2}\tau\cos 2\theta & |X\cos^{2}\tau\cos 2\theta|^{2} & X^{2}\cos^{4}\tau\cos 2\theta\sin 2\theta \\ (2 + X^{*}\cos^{2}\tau)X\cos^{2}\tau\sin 2\theta & X^{*2}\cos^{4}\tau\cos 2\theta\sin 2\theta & |X\cos^{2}\tau\sin 2\theta|^{2} \end{bmatrix}$$
(3.88)

We can then average this matrix over all possible angles θ and τ . The probability distributions p(..) for a random distribution are defined as shown in equation (3.89):

$$p(\theta) = \frac{d\theta}{2\pi} \quad -\pi \le \theta < \pi \qquad p(\tau) = \frac{\cos \tau d\tau}{2} \quad -\frac{\pi}{2} \le \tau < \frac{\pi}{2} \quad (3.89)$$

Clearly, in the face of such a distribution the off-diagonal elements of [T] will average to zero, as expected. However the three diagonal terms (the eigenvalues) maintain some scatterer information, as shown by explicit evaluation of the integrals in equation (3.91). The coherency matrix for such a volume may

be written in the form of an eigenvector decomposition, as shown in equation (3.90):

$$[T] = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$
(3.90)

where the eigenvalues take the explicit form shown in equation (3.91) (Boerner, 1992, Chapters I–4; Cloude, 1999; Ablitt, 2000).

$$\lambda_{1} = t_{11} = \frac{1}{4\pi} \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \int_{-\pi}^{\pi} (4 + 4X \cos^{2} \tau + X^{2} \cos^{4} \tau) \cos \tau d\theta d\tau$$

$$= 2 + \frac{4}{3}X + \frac{4}{15}X^{2}$$

$$\lambda_{2} = t_{22} = \frac{1}{4\pi} \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \int_{-\pi}^{\pi} X^{2} \cos^{4} \tau \cos^{2} 2\theta \cos \tau d\theta d\tau = \frac{2}{15}X^{2}$$

$$\lambda_{3} = t_{33} = \frac{1}{4\pi} \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \int_{-\pi}^{\pi} X^{2} \cos^{4} \tau \sin^{2} 2\theta \cos \tau d\theta d\tau = \frac{2}{15}X^{2}$$

(3.91)

In terms of A_p —the ratio of particle polarizabilities—the eigenvalues are given as shown in equation (3.92):

$$\lambda_1 = \frac{2}{15}(2A_P^2 + 6A_P + 7) \quad \lambda_2 = \frac{2}{15}(A_P - 1)^2 \quad \lambda_3 = \frac{2}{15}(A_P - 1)^2 \quad (3.92)$$

and hence the ratio of eigenvalues yields information about the shape and material composition of the particles in the volume through the parameter A_p , where $A_p \ge 0$ (see equation (3.77)), as shown in equation (3.93):

$$0 \le R_{\lambda} = \frac{\lambda_2}{\lambda_1} = \frac{A_P^2 - 2A_P + 1}{2A_P^2 + 6A_P + 7} \le \frac{1}{2} \implies A_P = \frac{(1 + 3R_{\lambda}) \pm \sqrt{5R_{\lambda}(3 - R_{\lambda})}}{1 - 2R_{\lambda}}$$
(3.93)

Figure 3.26 shows how the eigenvalue ratio R_{λ} changes with change in shape of parameter A_p . We notice three important points in the function. As A_p tends to zero, so the ratio tends to 1/7 = 0.1429, being the limit for a cloud of oblate spheroids. For $A_p = 1$ the ratio tends to zero and there is zero depolarisation (a cloud of spheres). As A_p tends to infinity we see that the ratio tends to a maximum value of 0.5 (dipole cloud), this being the most depolarising case.

If we seek to invert this relationship by measuring the eigenvalue ratio in order to estimate mean particle shape, then we face an ambiguity issue as shown in Figure 3.27. Here we see that for ratios less than 0.1429 we have both a prolate and oblate solution for the same ratio. For ratios greater than 0.1429 and less than 0.5, we obtain a unique prolate solution. For ratios between 0.5 and 1 we violate the assumptions of this model, and must consider other types of depolarisation or noise in the data.

A similar treatment can be used for scattering by a random cloud of chiral (or handed) particles. In this case the scattering vector for an individual particle



Fig. 3.26 Eigenvalue ratio of coherency matrix versus particle shape parameter $A_{\rm p}$

Fig. 3.27 Prolate (solid)/oblate (dash) ambiguity in particle shape estimation from coherency eigenvalue ratio

can be obtained from the scattering matrix (derived in equation (3.82)), and written for rotation θ , as shown in equation (3.94):

$$\underline{k} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos 2\theta & -\sin 2\theta \\ 0 & \sin 2\theta & \cos 2\theta \end{bmatrix} \begin{bmatrix} 2 + X \cos^2 \tau \\ X \cos^2 \tau \\ -2i\kappa X \cos^2 \tau \end{bmatrix}$$
$$= \begin{bmatrix} 2 + X \cos^2 \tau \\ X \cos^2 \tau (\cos 2\theta + 2i\kappa \sin 2\theta) \\ X \cos^2 \tau (\sin 2\theta - 2i\kappa \cos 2\theta) \end{bmatrix}$$
(3.94)

Working through a similar procedure as used for the non-chiral case, we find the following result for the eigen-decomposition of the coherency matrix for a chiral volume (Cloude, 2002b):

$$[T] = \frac{1}{2} \begin{bmatrix} \sqrt{2} & 0 & 0\\ 0 & 1 & i\\ 0 & i & 1 \end{bmatrix} \cdot \begin{bmatrix} \lambda_1 & 0 & 0\\ 0 & \lambda_2 & 0\\ 0 & 0 & \lambda_3 \end{bmatrix} \cdot \begin{bmatrix} \sqrt{2} & 0 & 0\\ 0 & 1 & -i\\ 0 & -i & 1 \end{bmatrix}$$
(3.95)

where the eigenvalues are now given by equation (3.96):

$$\lambda_{1} = \frac{2}{15} (2A_{P}^{2} + 6A_{P} + 7)$$

$$\lambda_{2} = \frac{2}{15} (1 + 4\kappa) (A_{P} - 1)^{2}$$

$$\lambda_{3} = \frac{2}{15} (1 - 4\kappa) (A_{P} - 1)^{2}$$
(3.96)

There are two important points raised by this result. Firstly, the eigenvectors are now complex; that is, they are no longer just the default Pauli scattering mechanisms. This highlights the importance of using the eigenvectors for a full characterization of the scattering process. For example, expanding equation (3.95) and using (3.96) we obtain the following non-diagonal form of the coherency matrix for volume scattering from a random cloud of chiral particles:

$$[T] = \frac{2}{15} \begin{bmatrix} 2A_P^2 + 6A_P + 7 & 0 & 0\\ 0 & (A_P - 1)^2 & -i4\kappa(A_P - 1)^2\\ 0 & i4\kappa(A_P - 1)^2 & (A_P - 1)^2 \end{bmatrix}$$
(3.97)

Secondly, the presence of chirality has caused a split in the smaller eigenvalues. Indeed, we note that the chirality parameter is directly related to the scattering anisotropy (see equation (2.79)), defined as the normalized difference of minor eigenvalues, as shown in equation (3.98):

$$A = \frac{\lambda_2 - \lambda_3}{\lambda_2 + \lambda_3} = 4\kappa \tag{3.98}$$

3.4.1 Volume scattering with reflection symmetry

In the azimuthal symmetry case considered above, we integrated over a random distribution of particle orientations. Now we consider generalization to consider an oriented volume; that is, volumes with a dominant orientation (taken as $\theta = 0^{\circ}$). To model this case, we first consider a uniform distribution of particle orientation angles centred around $\theta = 0^{\circ}$, as shown in equation (3.99):

$$p(\theta) = \frac{d\theta}{2\Delta} - \Delta \le \theta < \Delta \quad p(\tau) = \frac{\cos \tau d\tau}{2} - \frac{\pi}{2} \le \tau < \frac{\pi}{2} \quad (3.99)$$

where Δ now determines the width of the distribution ($\Delta = \pi$ reverting to the random case). In this case the coherency matrix for an individual particle, which has the form shown in equation (3.88), can be averaged to obtain a new

reflection symmetric form as shown in equation (3.100):

$$\langle [T] \rangle = \begin{bmatrix} t_{11} & t_{12} & 0 \\ t_{12}^* & t_{22} & 0 \\ 0 & 0 & t_{33} \end{bmatrix}$$

$$\rightarrow \begin{cases} t_{11} = 2 + \frac{4}{3}X + \frac{4}{15}X^2, & t_{12} = \frac{\sin(2\Delta)}{2\Delta} \left(\frac{2}{3}X + \frac{4}{15}X^2\right) \\ t_{22} = \frac{2}{15}X^2 \left(1 + \frac{\sin(4\Delta)}{4\Delta}\right), & t_{33} = \frac{2}{15}X^2 \left(1 - \frac{\sin(4\Delta)}{4\Delta}\right) \\ \end{cases}$$

$$(3.100)$$

The information contained in these equations becomes more apparent by plotting the data in the H/ α plane (see Section 2.4.2.4). We begin by considering the specific example of a dipole cloud (prolate spheroids with X >> 1). A single such particle has H = 0, $\alpha = 45^{\circ}$, and as shown in Figure 3.28, when we form a cloud of such particles in random orientations, so the entropy increases. However, the average scattering mechanism remains nearly constant around $\bar{\alpha} = 45^{\circ}$, permitting us to identify the mean particle shape independently of the angular spread of the particles in the cloud.

It is interesting to see if this useful property of the H/ α diagram extends to arbitrary particle shape. Figure 3.29 shows the results for both prolate and oblate particles (in steps of 2 dB in the ratio A_P and for distributions from 0 to π). Note that spheres always lie at the origin of the H/ α diagram. We note that the match is very good for prolate spheroids; the scattering mechanism changes only a small amount with entropy increase. The oblate zone is more restricted in the H/ α plane, and shows more clearly the depolarising nature of such clouds, even for a fixed particle orientation. This arises because in equation (3.100) we are still assuming a random distribution of particle tilt angles τ . We see, however, that



Fig. 3.28 Variation of entropy/alpha for a cloud of dipoles



Fig. 3.29 Entropy alpha variation for clouds of prolate and oblate particles

the prolate spheroids are much more sensitive to orientation effects, and display a much wider range of depolarisers. We identify two points of interest. The point P represents the most extreme depolariser: a random cloud of dipoles with H = 0.9464, $\overline{\alpha} = 45^{\circ}$. This point is often taken as a model for branch/needledominated forest scattering in radar remote sensing applications. For example, it is embedded in the Freeman–Durden decomposition approach (see Chapter 4). The point Q corresponds to a random cloud of oblate spheroids (H = 0.62, $\overline{\alpha} = 20^{\circ}$). This is more typical of leaf-dominated vegetation scattering, and shows less depolarisation than the prolate case.

3.4.2 Bistatic volume scattering

We can also use the solution for scattering by small spheroidal particles as an important example of bistatic volume scattering. Considering a single particle, the scattering matrix was derived in equation (3.81), and following integration over a random distribution we obtain the following average coherency matrix, where Ψ is the scattering angle ($\Psi = \pi$ corresponds to backscatter) and A_p is the particle shape function or ratio of polarizabilities.

$$\langle [T] \rangle = f \begin{bmatrix} t_{11} & t_{12} & 0 & 0 \\ t_{12}^* & t_{22} & 0 & 0 \\ 0 & 0 & t_{33} & 0 \\ 0 & 0 & 0 & t_{44} \end{bmatrix} \Rightarrow \begin{cases} t_{11} = z(1 - \cos\psi) + (1 + \cos\psi)^2 \\ t_{12} = (\cos^2\psi - 1) \\ t_{22} = z(1 + \cos\psi) + (1 + \cos\psi)^2 \\ t_{33} = z(1 + \cos\psi) \\ t_{44} = z(1 - \cos\psi) \end{cases}$$

$$f = \frac{2A_P^2 + 6A_P + 7}{30} \quad z = \frac{2(A_P - 1)^2}{2A_P^2 + 6A_P + 7}$$
(3.101)

The eigenvalues of the coherency matrix can then be calculated as shown in equation (3.102):

$$\lambda_{1,2} = fz + f(1 + \cos^2 \psi) \pm f \sqrt{(1 + \cos^2 \psi)^2 + z(z - 4)\cos^2 \psi}$$

$$\lambda_3 = fz(1 - \cos \psi)$$
(3.102)

$$\lambda_4 = fz(1 + \cos \psi)$$

(Note that the rank ordering of the eigenvalues depends on scattering angle.) For backscatter we note that $\lambda_4 = 0$ and the eigenvalues reduce to those shown in equation (3.96). We can use these equations to highlight two important examples. In the first we consider bistatic scattering from a cloud of spheres ($A_p = 1$). In this case, f = 0.5, z = 0, and the eigenvalues have the form shown in equation (3.103):

$$\lambda_1 = (1 + \cos^2 \psi), \lambda_2 = \lambda_3 = \lambda_4 = 0$$
 (3.103)

which shows that there is no depolarisation at any angle in single scattering from a cloud of spheres. (Note that for the moment we are ignoring multiple scattering contributions.) The scattering diagram with angle ψ is shown in Figure 3.30. The pattern arises from the combination of the uniform pattern for polarisation perpendicular to the page and the characteristic dumbbell shape for polarisation in the plane (dashed lines in Figure 3.30).

At the other extreme we can consider bistatic scattering from a cloud of small dipoles ($A_p >> 1$). In backscatter this yielded the strongest depolarisation of all single scattering configurations. In general bistatic scattering we obtain the



Fig. 3.30 Bistatic scattering diagram for a cloud of small spheres



Fig. 3.31 Bistatic variation of scattering eigenvalues for a dipole cloud

following form for the eigenvalues:

$$\lambda_{1,2} = \frac{A_p^2}{15} \left(2 + \cos^2 \psi \pm \sqrt{1 - \cos^2 \psi + \cos^4 \psi} \right)$$

$$\lambda_3 = \frac{A_p^2}{15} \left(1 - \cos \psi \right)$$

$$\lambda_4 = \frac{A_p^2}{15} \left(1 + \cos \psi \right)$$

(3.104)

Figure 3.31 shows how the eigenvalues (normalized so that $\Sigma \lambda = 1$) vary with scattering angle. Note how for backscatter and forward scattering the coherency matrix has rank 3, as expected from reciprocity. The other interesting point relates to 90-degree or lateral scattering. Here we see that the minor eigenvalues are all equal, corresponding to a noise subspace of the coherency matrix. Hence only at this angle can we give the scattering interpretation as a single [S] matrix (the dominant eigenvalue spectrum and hence structure to the depolarisation process.

It is useful to plot the scattering entropy variation corresponding to Figure 3.31. Figure 3.32 shows how bistatic entropy H varies with scattering angle. We note that nowhere is depolarisation complete (H = 1), and that for forward scatter and backscatter the entropy is smallest and hence depolarisation at a minimum. The reason for this is the loss of one eigenvalue to the reciprocity theorem, which, as we noted earlier, limits the level of depolarisation achievable in backscatter. The maximum depolarisation occurs for the lateral (90-degree) scattering case. This then represents the strongest depolariser so far encountered. However, we have considered only single scattering depolarisation. When waves encounter multiple scattering—when we account for



Fig. 3.32 Bistatic variation of scattering entropy for a dipole cloud

Fig. 3.33 Schematic representation of multiple scattering for backscatter problems

interactions between particles—then there exist possibilities for higher levels of depolarisation, as we now consider.

3.4.3 Depolarisation in multiple scattering

In the previous sections we have seen that there are limits imposed on the level of depolarisation generated by single scattering. In this section we briefly consider the more complicated problem of multiple scattering, and show how this can lead to much higher levels of scattering depolarisation (Macintosh, 1989; Bicout, 1992; Brosseau, 1994; Hovenier, 2004; Mishchenko, 2006). However, we highlight two important features of such phenomena. Firstly, how in backscatter the reciprocity theorem still acts, even in the presence of the most complicated multiple scattering processes, to limit the coherency matrix to rank 3 and hence maintain a cap on the level of depolarisation that can occur. The second point we consider is how depolarisation increases with increasing order of multiple scattering, to highlight some interesting anisotropies that occur in the variation of depolarisation (van Albada, 1988; Mishchenko, 2006; Macintosh 1989).

The process of multiple scattering is shown schematically in Figure 3.33. Here we show an incident field interacting first with particle 1, the scattered radiation from which then interacts with particle 2, and so on, up to order n, before being scattered in the direction of the receiver (here backscatter).

Hence the coherency matrix for such a scenario is not simply given by the sum of coherency matrices for the n individual particles, but involves mutual interactions. These interactions give rise to additional sources of depolarisation, as we now demonstrate.

3.4.3.1 The Mishchenko decomposition

While a detailed analysis of all these multiple interactions is complicated, there are three groupings that tend to dominate the response. The first are just the summed direct single scattered returns from each particle. The second are the ladder terms in the expansion of the multiple scattering integral equations (Mishchenko, 1992). These arise from all multiple paths from 1-2-3 and so on, and provide a new source of depolarisation. It happens that these ladder terms are relatively easy to calculate using the techniques of vector radiative transfer theory (VRT) (Jin, 1994b; Hovenier, 2004). This approach is based on a set of integral equations set up as a balance of energy loss due to propagation, extinction and scattering, and yields a set of linear equations that can be solved numerically or even analytically for some simple cases (Tsang, 1985; Hovenier, 2004). There is, however, a third important contribution not included even in these ladder terms. These are the cyclical components of the integral equation, and basically these combine those multiple paths through the medium which have a high coherence between a path and its time-reversed form, as shown on the left and right of Figure 3.33. In general, therefore, we can write the Mueller and coherency matrices for multiple scattering problems as a sum of three terms, called the Mishchenko decomposition-first derived in Mishchenko (1992)—as shown in equation (3.105):

$$\langle [M] \rangle = \langle [M_S] \rangle + \langle [M_L] \rangle + \langle [M_C] \rangle \Leftrightarrow \langle [T] \rangle = \langle [T_S] \rangle + \langle [T_L] \rangle + \langle [T_C] \rangle$$
(3.105)

where the subscripted S is for single scattering, L for the ladder terms, and C for the cyclical terms. Concentrating on the new polarisation effects caused by multiple scattering and their effect on the eigenvalue spectrum of [T], we further rewrite the coherency matrix in the form shown in equation (3.106), where 'MS' now contains all the multiple scattering contributions:

$$\langle [T] \rangle = \langle [T_S] \rangle + \langle [T_L] \rangle + \langle [T_C] \rangle = \langle [T_S] \rangle + \langle [T_{MS}] \rangle$$
(3.106)

We can shed some light on the structure of these new depolarising terms by considering the multiple paths in Figure 3.33 and writing the multiple scattering terms as a function of their scattering vectors, as shown in equation (3.107):

$$\langle [T_{MS}] \rangle = \left\langle \left(\underline{k}_{(1,n)} + \underline{k}_{(n,1)} \right) \left(\underline{k}_{(1,n)} + \underline{k}_{(n,1)} \right)^{*T} \right\rangle$$

$$= \left\langle \underline{k}_{(1,n)} \underline{k}_{(1,n)}^{*T} + \underline{k}_{(n,1)} \underline{k}_{(n,1)}^{*T} + \underline{k}_{(1,n)} \underline{k}_{(n,1)}^{*T} + \underline{k}_{(n,1)} \underline{k}_{(1,n)}^{*T} \right\rangle$$

$$= \left\langle \underline{k}_{(1,n)} \underline{k}_{(1,n)}^{*T} + \underline{k}_{(n,1)} \underline{k}_{(n,1)}^{*T} \right\rangle + \left\langle \underline{k}_{(1,n)} \underline{k}_{(n,1)}^{*T} + \underline{k}_{(n,1)} \underline{k}_{(1,n)}^{*T} \right\rangle$$

$$= \left\langle [T_L] \right\rangle + \left\langle [T_C] \right\rangle$$

$$(3.107)$$

We see that the ladder terms are formed from averages over conventional coherency matrix factors. However, the cyclical terms involve a mixture of path and time-reversed path contributions. The calculation of these latter terms is particularly difficult, but there is an interesting relationship between ladder and cyclical terms in the *exact* backscatter direction. This link arises because of the reciprocity theorem, which relates time-reversed paths as an apparent exchange of transmitter and receiver positions. We have already seen in equation (1.147) that the two coherent [S] matrices and corresponding scattering vectors for the two paths are related as shown in equation (3.108):

$$\begin{bmatrix} S_{(1,n)} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} S_{(n,1)} \end{bmatrix}^T \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \Rightarrow \underline{k}_{(1,n)} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \underline{k}_{(n,1)}$$
(3.108)

When we substitute this connection into equation (3.107) we see that if we know the contribution of the ladder terms (from VRT, for example) then the cyclical terms can be immediately estimated as shown in equation (3.109):

$$\langle [T_L] \rangle = \begin{bmatrix} t_{11} & t_{12} & 0 & t_{14} \\ t_{12}^* & t_{22} & 0 & t_{24} \\ 0 & 0 & t_{33} & 0 \\ t_{14}^* & t_{24}^* & 0 & t_{44} \end{bmatrix} \Rightarrow \langle [T_C] \rangle = \begin{bmatrix} t_{11} & t_{12} & 0 & t_{14} \\ t_{12}^* & t_{22} & 0 & t_{24} \\ 0 & 0 & -t_{33} & 0 \\ t_{14}^* & t_{24}^* & 0 & t_{44} \end{bmatrix}$$
(3.109)

Note that this means that the ladder terms by themselves violate the reciprocity theorem, and care must be exercised when using predictions from VRT for exact backscatter calculations from random media. It is only when we add the 'L' and 'C' terms that we obtain a rank-3 coherency matrix in equation (3.106).

Turning now to an example of depolarisation by multiple scattering in a direction other than backscatter, we consider forward incoherent scattering by a random cloud of particles. From symmetry arguments, the Mueller and coherency matrices for such a scenario must have the normalized form shown in equation (3.110):

$$[M] = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \delta & 0 & 0 \\ 0 & 0 & \delta & 0 \\ 0 & 0 & 0 & \varepsilon \end{bmatrix} \Leftrightarrow [T] = \begin{bmatrix} 1+\varepsilon+2\delta & 0 & 0 & 0 \\ 0 & 1-\varepsilon & 0 & 0 \\ 0 & 0 & 1-\varepsilon & 0 \\ 0 & 0 & 0 & 1+\varepsilon-2\delta \end{bmatrix}$$
$$\Rightarrow \begin{cases} |\varepsilon| \le 1 \\ |\delta| \le 1 \\ 2\delta - \varepsilon \le 1 \end{cases}$$
(3.110)

We see that the diagonal elements must obey some restrictions in order that the eigenvalue spectrum of [T] be non-negative, and that the rotational subspace eigenvalues be equal. We saw above that single scattering from such a cloud would yield a rank-1 coherency matrix and zero depolarisation ($\varepsilon = \delta = 1$). However, conditions change when we allow multiple scattering. This problem, when formulated for a cloud of spheres, can be determined analytically using VRT (at least for the ladder terms contributions; and here we are assuming that the coherent forward scattered wave is completely attenuated). It can be expressed as a function of the order of multiple scattering n, with details in

Brosseau, 1994), and here we state the main result, shown in equation (3.111). We show both the Mueller and corresponding coherency matrix forms of the result as a function of n. (Note that in this case there is no dependence on bistatic scattering angle, as we are restricting attention to forward scattering.) For example, for n = 0 (single scattering) we obtain a Mueller matrix with the form of the 4 × 4 identity matrix, and a rank-1 coherency matrix with three zero eigenvalues. These correspond to forward scattering by a sphere matrix.

$$\langle [M] \rangle = \begin{bmatrix} 1 + \frac{1}{2} \left(\frac{7}{10}\right)^n & 0 & 0 & 0 \\ 0 & \frac{3}{2} \left(\frac{7}{10}\right)^n & 0 & 0 \\ 0 & 0 & \frac{3}{2} \left(\frac{7}{10}\right)^n & 0 \\ 0 & 0 & 0 & \frac{3}{2} \left(\frac{1}{2}\right)^n \end{bmatrix} \Rightarrow \langle [T] \rangle$$

$$= \frac{3}{2} \begin{bmatrix} \frac{2}{3} + \left(\frac{7}{10}\right)^n \left(\frac{7}{3} + \left(\frac{5}{7}\right)^n\right) & 0 & 0 & 0 \\ 0 & \frac{2}{3} + \left(\frac{7}{10}\right)^n \left(\frac{1}{3} - \left(\frac{5}{7}\right)^n\right) & 0 & 0 \\ 0 & 0 & \frac{2}{3} + \left(\frac{7}{10}\right)^n \left(\frac{1}{3} - \left(\frac{5}{7}\right)^n\right) & 0 \\ 0 & 0 & 0 & \frac{2}{3} + \left(\frac{7}{10}\right)^n \left(\frac{1}{3} - \left(\frac{5}{7}\right)^n\right) & 0 \\ 0 & 0 & 0 & \frac{2}{3} + \left(\frac{7}{10}\right)^n \left(\frac{1}{3} - \left(\frac{5}{7}\right)^n\right) & 0 \\ (3.111) \end{cases}$$

Figure 3.34 shows how the eigenvalues of this coherency matrix vary with scattering order *n*. Note that for n = 0 (single scattering) the matrix has only one non-zero eigenvalue, as expected, while as *n* increases so the normalized eigenvalues all tend to 0.25; that is, to equality and maximum depolarisation (H = 1). Figure 3.35 shows the corresponding multiple scattering entropy as a function of increasing order. Note that the entropy approaches 1—a perfect depolariser for orders n > 5 (corresponding to $\varepsilon = \delta = 0$ in equation (3.110)). A second interesting observation concerns the anisotropy of the eigenvalues. Two of the minor eigenvalues are always equal, as expected from symmetry,



Fig. 3.34 Variation of forward scattering eigenvalues with scattering order *n*



Fig. 3.35 Forward scattering entropy and anisotropy versus scattering order *n*

and so there is only one anisotropy, A_{34} , to consider. The variation of this anisotropy with order is also shown in Figure 3.35.

Note how the anisotropy is high for low orders of multiple scattering (n > 0) and decreases steadily with increasing order. This implies that although the spheres depolarise, they do not do so equally for all incident polarisations. We see from the form of the Mueller matrix that this anisotropy is due to $\varepsilon < \delta$ for all scattering orders. This means that the multiple scattering depolarises incident circular polarisations, with a Stokes vector of the form $(1, 0, 0, \pm 1)$ more than linear, with Stokes vector $(1, \pm 1, 0, 0)$, with the difference between the two decreasing with increasing scattering order.

3.5 Simple physical models for volume scattering and propagation

We now turn to a different aspect of volume backscattering: namely, how to express the *total* scattering by a cloud of particles as a simplified volume integral. The motivation for doing this is to enable inversion of these models from scattered field data. In the most general case we can formulate vector scattering in random media using vector radiative transfer (VRT) theory (Chandrasekhar, 1960; Jin, 1994; Iniesta, 2003; Hovenier, 2004). In simple geometries such as planar stratified media, there now exist efficient methods for solving these equations to include both single and multiple scattering effects. However, resort must often be made to advanced numerical techniques. For this reason, low-order simplified solutions are still useful in obtaining a basic physical understanding and to enable model inversion via parameter estimation. We illustrate this by employing two such examples: the water cloud model (WCM) (Attema, 1978; Ulaby, 1986), which forms the basis for many practical algorithms in radar remote sensing; and a two-phase anisotropic vegetation model from Ulaby


Fig. 3.36 Geometry of the water cloud model (WCM)

(1986) that can be used to demonstrate the importance of oriented volume effects in wave propagation through clouds of anisotropic particles.

3.5.1 The water cloud model (WCM)

To illustrate how volume scattering calculations typically proceed from the solution for an individual particle into a cloud or scattering ensemble, we consider a simple but important example called the water cloud model (WCM), first derived in Attema (1978), which represents a volume as a collection of identical particles in random positions, with density N_v particles per unit volume. For simplicity, each particle is assumed spherical with radius r_i , and furthermore we ignore multiple scattering between particles. The scattering geometry is shown in Figure 3.36. A plane wave of power density P Wm⁻² is incident at angle θ , and we seek an expression for the backscatter coefficient $\sigma_0(m^2m^{-2})$ as a function of the volume properties, defined as follows:

 N_v (m⁻³): number of particles per unit volume *h* (m): height of layer θ (rad): angle of incidence

Under the assumption of independent scattering from each particle, and ignoring any multiple scattering effects, we can then postulate the following models for the volume backscattering coefficient σ_v (m²m⁻³) and extinction coefficient σ_e (m⁻¹) of the medium, as shown in equation (3.112):

$$\sigma_{\nu} = \sum_{i=1}^{N_{\nu}} \sigma_{Pi} \quad \sigma_e = \sum_{i=1}^{N_{\nu}} Q_{Pi} \tag{3.112}$$

Here, σ_p and Q_p are the single particle backscattering and extinction crosssections respectively, defined as the ratios of scattered and absorbed powers to the incident power density, and both having units of m². For Rayleigh scattering, when the particles are small compared to a wavelength these parameters can be written explicitly in terms of the particle size and material composition, as shown in equation (3.113):

$$Q_{P} = Q_{A} + Q_{S} = \frac{\lambda^{2}}{\pi} \operatorname{Im}\left(\frac{1-\varepsilon_{r}}{\varepsilon_{r}+2}\right) (\beta_{b}r)^{3} + \frac{2\lambda^{2}}{3\pi} \left|\frac{\varepsilon_{r}-1}{\varepsilon_{r}+2}\right| (\beta_{b}r)^{6}$$

$$\approx Q_{A} = \frac{8\pi^{2}}{\lambda_{o}} \operatorname{Im}\left(\frac{1-\varepsilon_{r}}{\varepsilon_{r}+2}\right) r^{3}$$

$$\sigma_{P} = \frac{64\pi^{5}}{\lambda_{o}^{4}} \left|\frac{\varepsilon_{r}-1}{\varepsilon_{r}+2}\right| r^{6}$$
(3.113)

Note the following two points:

- 1 The extinction is composed of two parts: an absorption (A) and a scattering (S) loss. For Rayleigh scattering (with $\beta_b r$ small), as shown, the absorption loss is much higher than the scattering loss, and so the latter is often ignored.
- 2 The important scaling parameter is the product $\beta_b r$, where β_b is the wavenumber in the background medium $\beta_b = \frac{2\pi}{\lambda_a} \sqrt{\varepsilon'_{rb}}$, and λ_o is the

wavelength in free space. If we assume that the background is just air $(\varepsilon'_{rb} = 1)$ then we can express Q_P in terms of particle radius and complex dielectric constant ε_r , as shown in equation (3.113).

With this in place we can express the backscattering coefficient as an integral over all the particles, as shown in equation (3.114):

$$P_{i} = PA \cos \theta$$

$$P_{r} = \frac{PA}{4\pi R^{2}} \cos \theta \int_{0}^{h/\cos \theta} \sigma_{v} \exp(-2\sigma_{e}z) dz$$

$$\Rightarrow \sigma_{o}^{WCM}(\theta) = 4\pi R^{2} \frac{P_{r}}{P_{i}} = \frac{\sigma_{v} \cos \theta}{2\sigma_{e}} (1 - \exp(-2\sigma_{e}h \sec \theta))$$

$$\Rightarrow \sigma_{o}^{WCM}(\theta) = \frac{\sigma_{v} \cos \theta}{2\sigma_{e}} \left(1 - \frac{1}{L^{2}(\theta)}\right) \quad L(\theta) = e^{\sigma_{e}h \sec \theta}$$
(3.114)

Here we see that the scattering cross-section can be expressed in terms of a combination of propagation (σ_e) and scattering (σ_v), with the former appearing in two ways: in the ratio $\frac{\sigma_v}{2\sigma}$, and in the one-way loss factor L(θ).

The key idea of the water cloud model is to realize that the dominant material influence in many microwave volume scattering problems comes from the water content m_v of the volume, measured in gm^{-3} . Hence the idea is to consider the volume as a cloud of spherical water particles and recalculate the backscatter in terms of water content. To do this we need two key results. The first concerns the relation between total extinction and water content. The starting point is to express the water content as the product of water density (10^6 gm^{-3}) and the fraction of space occupied by the spherical particles, which yields the following linear relation involving constant C_1 between extinction and water content:

$$m_{\nu} = 10^{6} \sum_{i=1}^{N_{\nu}} \frac{4\pi}{3} r_{i}^{3}$$

$$\sigma_{e} = \frac{8\pi^{2}}{\lambda_{o}} \operatorname{Im}\left(\frac{1-\varepsilon_{r}}{\varepsilon_{r}+2}\right) \sum_{i=1}^{N_{\nu}} r_{i}^{3} \end{cases} \Rightarrow \sigma_{e} = \frac{6\pi \times 10^{-6}}{\lambda_{o}} \operatorname{Im}\left(\frac{1-\varepsilon_{r}}{\varepsilon_{r}+2}\right) m_{\nu} = C_{1}m_{\nu}$$
(3.115)

The second key idea in the WCM is that the ratio $\frac{\sigma_v}{2\sigma_e}$ is often independent of water content. This follows from consideration of the simplest case, when we have a cloud of identical particles of the same radius $r_i = r$, and when we can write the ratio as a constant C_2 , as shown in equation (3.116):

$$\left. \begin{array}{l} \sigma_{v} = \sum_{i=1}^{N_{v}} \sigma_{Pi} = N_{v} \sigma_{P} \\ \sigma_{e} = \sum_{i=1}^{N_{v}} Q_{Pi} = N_{v} Q_{P} \end{array} \right\} \Rightarrow \frac{\sigma_{v}}{2\sigma_{e}} = \frac{\sigma_{P}}{2Q_{P}} = C_{2}$$
(3.116)

Note that the assumption of identical particle size can be relaxed by including a new function—the particle size distribution function p(r)—so that the expression for total extinction (and other parameters) takes on an integral form as

shown in equation (3.117):

$$\sigma_e = \int_{r_1}^{r_2} p(r)Q_e(r)dr$$
 (3.117)

While no simple forms exist for the size distribution in vegetation scattering, for example, results are often taken from atmospheric cloud physics, where drop size distributions are better characterized (Ulaby, 1986). The main consequence of allowing a particle size distribution is that the backscatter σ_v becomes a quadratic function of m_v . Thus, in this more complicated version of the WCM, we can combine this with the linear relation between extinction and m_v to postulate a linear relation between the ratio and m_v , as shown in equation (3.118):

$$\frac{\sigma_{\nu}}{2\sigma_{e}} \approx C_{3}m_{\nu} \tag{3.118}$$

Finally, by combining all of these ideas we can then rewrite the expression for backscatter from a volume in terms of the water content and angle of incidence, as shown in equation (3.119), for both the uniform particle assumption (3.119a) and the distributed particle size case (3.119b):

$$\sigma_o^{WCM}(\theta) = C_2 \cos \theta [1 - \exp(-C_1 m_v h \sec \theta)]$$
(3.119a)

$$\sigma_o^{WCM}(\theta) = C_3 m_v \cos \theta [1 - \exp(-C_1 m_v h \sec \theta)]$$
(3.119b)

This example illustrates how we can often formulate volume scattering problems in terms of a small number of physical parameters, and then use scattering models to simplify the expressions. However, we note that from a polarisation point of view the predicted behaviour of the WCM is trivial. With the assumption of independent identical spherical particles and single scattering, the scattering matrix for each is just the 2×2 identity, and hence there is zero cross- or depolarisation prediction and equal scattering (in amplitude and phase) for the copolar channels. In the next section we shall see how to extend this model to account for non-spherical particles. We shall also make use of this model again in a coherent form when considering polarimetric interferometry in Chapter 7. In order to illustrate how polarisation effects can be included in volume scattering we turn now to consider a more complicated version of the WCM.

3.5.2 Ulaby model: two-phase volume propagation models

A host medium (with background dielectric constant ε_h) with one type of particle inclusion is called a two-phase mixture. The effective dielectric constant of the composite medium can then be expressed in terms of three parameters: ε_h , the complex dielectric constant of the included particles ε_I and their volume fraction v_i , defined for ellipsoidal particles as shown in equation (3.120), where *a*, *b* and *c* are the particle dimensions, and N_v is the number of ellipsoids per



Fig. 3.37 Three examples of volume scattering: prolate (left), spheres (centre), and oblate (right)

unit volume (Ulaby, 1986, Appendix E-4).

$$v_i = \frac{4}{3}\pi abcN_v \tag{3.120}$$

In most applications of interest (microwave vegetation and forest scattering, for example), the volume fraction is very small (< 0.01). In what follows we restrict attention to models applicable in this small-volume fraction limit. We also consider a low-frequency approach that requires the particle diameter to be less than a wavelength in the particle material; that is, for spheroids, $a = b << \lambda$. The dielectric constant of such a composite medium is then given by a mixture model, such as the Polder–Van Santen/de Loor formula, shown in equation (3.121):

$$\varepsilon_r = \varepsilon_h + \frac{v_i(\varepsilon_i - \varepsilon_h)}{1 + L_u(\frac{\varepsilon_i}{\varepsilon_h} - 1)}$$
(3.121)

Here the parameter L_u is the particle shape factor of the ellipsoid along its u axis, as already encountered in equation (3.76). We now consider three types of volume of interest: one with embedded needles or prolate spheroids as shown on the left side of Figure 3.37; one with spheres, at the centre of the figure; and one with discs (oblate spheroids) on the right of the figure. In all cases the positions of the particles are assumed random, but their orientations are the same. For each of these we can write the following special forms of equation (3.121) for the effective dielectric constant of the medium.

a) Case 1: A volume of needles aligned along the z axis ($L_a = L_b = 0.5, L_c = 0$). In this case we have different dielectric constants for different polarisations. For the electric field polarised in the x, y and z directions we have the following values:

$$\varepsilon_r^x = \varepsilon_h \left[1 + 2v_i \frac{(\varepsilon_i - \varepsilon_h)}{(\varepsilon_i + \varepsilon_h)} \right]$$

$$\varepsilon_r^y = \varepsilon_r^x$$

$$\varepsilon_r^z = \varepsilon_h + v_i (\varepsilon_i - \varepsilon_h)$$

(3.122)

b) Case 2: A volume of spheres (L = 1/3). In this case the dielectric constant is independent of polarisation, and is given by the following expression:

$$\varepsilon_r = \varepsilon_h \left[1 + 3\nu_i \frac{(\varepsilon_i - \varepsilon_h)}{(\varepsilon_i + 2\varepsilon_h)} \right]$$
(3.123)

c) Case 3: A volume of discs oriented in the x-y plane ($L_a = L_b = 0, L_c = 1$). Here again we have anisotropic behaviour, with the dielectric constant being polarisation dependent, as shown in equation (3.124):

$$\varepsilon_r^x = \varepsilon_h + v_i(\varepsilon_i - \varepsilon_h)$$

$$\varepsilon_r^y = \varepsilon_r^x$$

$$\varepsilon_r^z = \varepsilon_h \left[1 + v_i \left(1 - \frac{\varepsilon_h}{\varepsilon_i} \right) \right]$$
(3.124)

The above examples are all for so-called oriented volumes, where the particles are aligned to form a crystal-type structure. However, in many practical applications the particles show no particular alignment (forming a random volume). For random orientation the dielectric constant no longer depends on polarisation, and the scalar mixture formula takes the following form:

$$\varepsilon_r = \varepsilon_h + \frac{v_i}{3} (\varepsilon_i - \varepsilon_h) \sum_{u=a,b,c} \frac{1}{1 + A_u(\frac{\varepsilon_i}{\varepsilon_h} - 1)}$$
(3.125)

It can be seen that this involves an average over the shape functions along the three principle particle axes. For a cloud of spheres such averaging, of course, makes no difference, and the effective medium permittivity is the same as equation (3.123). For the two other cases, however, we obtain the following modified results.

d) Case 4 : For a random cloud of oblate spheroids (discs):

$$\varepsilon_r = \varepsilon_h + \frac{v_i}{3}(\varepsilon_i - \varepsilon_h) \left(2 + \frac{\varepsilon_h}{\varepsilon_i}\right)$$
 (3.126)

e) Case 5 : For a random cloud of prolate spheroids (needles):

$$\varepsilon_r = \varepsilon_h + \frac{v_i(\varepsilon_i - \varepsilon_h)(5\varepsilon_h + \varepsilon_i)}{3(\varepsilon_h + \varepsilon_i)}$$
(3.127)

We can now make use of these results to estimate the polarisation dependent oneway loss factor $L(\underline{w}, \theta)$ for a slab of vegetation. Figure 3.38 shows a schematic of the problem. A wave is incident at θ_0 from the normal, and traverses a slab of thickness h. In particular, we consider the slab to be constructed of three phases of material, the first—a host background—being free space, so that $\varepsilon_h = 1$. The second phase is a random collection of discs, modelling the leafy or fine-scale structure of the vegetation. Each leaf has a complex dielectric constant ε_L . Critically, for this phase the dielectric constant is independent of polarisation. Finally, we model the stalks or trunks as a vertically oriented array of prolate spheroids with complex particle dielectric constant ε_s , as shown schematically in Figure 3.39. These will have a higher dielectric constant for vertical polarisation than for horizontal.

The loss factor for the composite medium can then be written as the product of polarisation independent (for the random component) and dependent (for the oriented volume) terms as shown in equation (3.128):

$$L\left(\underline{w}\right) = L_L L_S\left(\underline{w}\right) \tag{3.128}$$

The random component is given by an exponential propagation factor, as shown in equation (3.129), where n = n' - in'' is the complex refractive index of the



Fig. 3.38 Geometry of wave propagation through a slab of vegetation



Fig. 3.39 Composite vegetation layer composed of leaf and stalk contributions

two-phase material composed of leaf + background.

$$L_L = e^{-\frac{2\pi h}{\cos\theta_0 \lambda} n_L''} \tag{3.129}$$

For small volume fractions this refractive index can be related to the dielectric constant as shown in equation (3.130):

$$n_L'' = \left| \operatorname{Im}(\sqrt{\varepsilon_r} \right| \approx \frac{\varepsilon_r''}{2} \Rightarrow L_L = e^{-\frac{\pi h \varepsilon_r''}{\cos \theta_0 \lambda}}$$
(3.130)

Finally, we use Case 4 from above, setting $\varepsilon_h = 1$ and $\varepsilon_L'' >> 1$ (assuming the leaves have a high water content) to obtain an expression for the one-way loss factor as a function of the particle volume fraction and imaginary part of dielectric constant, as shown in equation (3.131).

$$\varepsilon_r'' \approx \frac{2\nu_i \varepsilon_L''}{3} \Rightarrow L_L = e^{-\frac{2\pi}{3\lambda \cos\theta_0} h \nu_L \varepsilon_L''}$$
(3.131)

An important practical measure of the leaf coverage is the leaf-area index, or LAI, defined as the total single-sided area of all the leaves over a unit area of ground. From this definition we can then rewrite the loss factor in terms of LAI as shown in equation (3.132), where t_L is the thickness of the leaf particles.

$$LAI = \frac{hv_L}{t_L} \Rightarrow L_L = e^{-\frac{2\pi}{3\lambda\cos\theta_0}t_L LAI\varepsilon_L''}$$
(3.132)

The oriented stem loss factors cause a differential extinction; that is, a difference between the eigenpolarisations, which in this case are H and V polarisation. The HH polarisation is the easiest to consider first (polarisation perpendicular to the plane of the page in Figure 3.38). This has a dielectric constant expressed solely in terms of the stalk volume fraction (again using the approximation that $|\varepsilon_s| \gg 1$) as shown in equation (3.133):

$$\varepsilon_r^y = 1 + 2\nu_s \frac{(\varepsilon_s - 1)}{(\varepsilon_s + 1)} = (1 + 2\nu_s) - i\left(\frac{4\nu_s \varepsilon_s''}{(\varepsilon_s'')^2 + (\varepsilon_s')^2}\right)$$
(3.133)

This shows that the H polarised wave is only weakly influenced by the presence of the stalk components, and the extinction in this channel will be dominated by the leafy component given in equation (3.126) (which affects all polarisations equally).

The VV polarisation can be obtained by projecting V onto the x and z axes and then applying the appropriate component from Case 1 (equation 3.122). For example, Figure 3.38 shows how to calculate the z component of the electric field from a geometrical factor $\sin \theta_0$. Likewise, the x component is given by $\cos \theta_0$ and this component 'sees' the same dielectric constant as HH. The vertical component sees a different extinction given from Case 1, as shown in equation (3.134):

$$\varepsilon_r^z \approx \left(1 + v_s \left(\varepsilon_s' - 1\right)\right) - i v_s \varepsilon_s''$$
 (3.134)

In this way the loss factors for HH and VV polarisation due to the stalk components can be written as shown in equation (3.135). Note that for normal incidence the effect of the vertical stalks is minimal and the extinction is isotropic (the same in HH and VV). Only as the angle of incidence increases, so the differential effect increases, with the stalks having maximum influence at 90° or grazing incidence.

$$L_{s}(\underline{w}) \Rightarrow \begin{cases} L_{HH} \approx \exp\left(-\frac{4\pi n_{HH}'' h}{\lambda \cos \theta_{o}}\right) & n_{HH}'' = \operatorname{Im}\left(\sqrt{\varepsilon_{r}^{y}}\right) \\ L_{VV} \approx \exp\left(-\frac{4\pi \left(\cos^{2} \theta_{o} n_{HH}'' + \sin^{2} \theta_{o} n_{VV}''\right) h}{\lambda \cos \theta_{o}}\right) & n_{VV}'' = \operatorname{Im}\left(\sqrt{\varepsilon_{r}^{z}}\right) \end{cases}$$
(3.135)

This model shows how extinction propagation through oriented volumes can be related to simple physical parameters such as volume fractions and dielectric constants. The above model, however, is restricted to low-frequency scattering and in particular to low albedo problems (when the ratio of scattering loss to total extinction is small).

3.5.3 Forest extinction models

The extension of these models to more general cases is complicated by the need to consider both scattering and absorption contributions to the extinction (Tsang, 1985). One way to deal with such complexity is to employ empirical models generated by fitting functions (usually polynomials) to experimental data. For example, in forestry, using a comprehensive set of airborne data across different forest sites and using different airborne sensors (including the P-3 (UHF), Carabas II (VHF), JPL-AIRSAR (P,L,C), and SRI(UHF)), the following model has been proposed to represent the total two-way extinction *X* in dB (Bessette, 2001):

$$X\cos\theta = a.F^b \tag{3.136}$$

where *F* is the frequency in MHz and *a* and *b* are regression parameters obtained from a global fit across all datasets. The mean values of these two parameters for 20, 50 and 80% of data are shown in Table 3.1. Figure 3.40 shows plots of extinction at the 20, 50 and 80% levels versus frequency for normal incidence $(\theta = 0^\circ)$ over the range 100 MHz to 1.3 GHz. Figure 3.41 shows the accompanying multiplicative scale factor for increasing angle of incidence. To find the total two-way extinction at frequency *F* for a given angle of incidence θ ,

 Table 3.1 Regression-based extinction model parameters (from Bessette (2001))

	20%	20%	50%	50%	80%	80%
Attenuation factors	a	b	a	b	a	b
HH	0.08	0.59	0.18	0.53	0.19	0.56
VV	0.21	0.47	0.3	0.47	0.32	0.50



Fig. 3.40 Example of an empirical model for predicting extinction through forest (dash = VV; solid = HH)

Fig. 3.41 Scale factor to account for increased extinction at larger angles of incidence

take the value from Figure 3.40 and multiply its dB value by the scale factor obtained from Figure 3.41. We note the following general points:

- 1. The extinction increases with frequency at a rate around $(6/\cos\theta)$ dB/decade for the 80% coverage fit.
- 2. The extinction for VV (shown as the dashed line in Figure 3.40) is always greater than for HH, indicating an oriented volume effect, even in complex forestry, with a slight preference for vertically oriented scatterers in the volume.

3.5.4 Dual polarised surface and volume depolarisation

Active remote sensing systems often operate in dual polarisation or compact modes (see Chapter 9) (Raney, 2006; Souyris, 2005). In this case, rather than measurement of the full [S] matrix, the system obtains a projection formed by using only a single transmit polarisation, typically either just horizontal H or vertical V and dual channel (H and V) coherent receive. These systems can still be used to partially characterize depolarisation effects by random surface and volume scatterers by employing 2×2 coherency matrixes (N = 2 depolarisation) as follows. For example, Bragg surface scattering from smooth surfaces will have 2×2 coherency matrices of the form shown in equation (3.137):

$$[J_H]_{Bragg} = \frac{1}{2} \begin{bmatrix} a + c + 2\operatorname{Re}(b) & 0\\ 0 & 0 \end{bmatrix}$$

$$[J_V]_{Bragg} = \frac{1}{2} \begin{bmatrix} a + c - 2\operatorname{Re}(b) & 0\\ 0 & 0 \end{bmatrix}$$
(3.137)

Note that these matrices imply zero scattering entropy and $\alpha_2 = 0$ for *all* angles of incidence and dielectric constants. We can extend this to include depolarisation caused by surface roughness by considering the X-Bragg model. This takes the projected form shown in equation (3.138):

$$[J_H]_{XBragg} = \frac{1}{2} \begin{bmatrix} a + \frac{c}{2} \left(1 + \frac{\sin 4\Delta}{4\Delta} \right) + 2\operatorname{Re}(b) \frac{\sin 4\Delta}{4\Delta} & 0 \\ 0 & \frac{c}{2} \left(1 - \frac{\sin 4\Delta}{4\Delta} \right) \end{bmatrix}$$
$$[J_V]_{XBragg} = \frac{1}{2} \begin{bmatrix} a + \frac{c}{2} \left(1 + \frac{\sin 4\Delta}{4\Delta} \right) - 2\operatorname{Re}(b) \frac{\sin 4\Delta}{4\Delta} & 0 \\ 0 & \frac{c}{2} \left(1 - \frac{\sin 4\Delta}{4\Delta} \right) \end{bmatrix}$$
(3.138)

Here we see non-zero scattering entropy, but note a mixture of roughness and dielectric constant dependence in the terms. Contrast this with the full coherency matrix formalism where we are able to separate roughness and moisture effects. Still, the level of depolarisation in equation (3.138) is small, and to see how high the entropy can be we turn again to the case of volume scattering from a random cloud of anisotropic particles. Using the notation of equation (3.92) we see that the dual polarisation coherency matrices can be expressed in terms of the particle anisotropy A_p , as shown in equation (3.139):

$$[J_H]_{Vol} = [J_V]_{Vol} = \frac{1}{15} \begin{bmatrix} 3A_p^2 + 4A_p + 8 & 0\\ 0 & (A_p - 1)^2 \end{bmatrix}$$
(3.139)

For dipole scatterers, A_p tends to infinity and we have the strongest depolariser. For oblate particles, on the other hand, $A_p = 0$ and the depolarisation is weaker. These two cases may be distinguished in dual polarised systems using their limiting form of 2×2 coherency matrix, as shown in equation (3.140):

$$[J_H]_{prolate} = [J_H]_{prolate} \propto \begin{bmatrix} 3 & 0 \\ 0 & 1 \end{bmatrix} \Rightarrow H_2 = 0.811$$

$$[J_H]_{oblate} = [J_V]_{oblate} \propto \begin{bmatrix} 8 & 0 \\ 0 & 1 \end{bmatrix} \Rightarrow H_2 = 0.503$$

(3.140)

We see that the maximum dual polarised (N = 2) entropy from such a cloud is 0.811, and note that while dual polarised systems offer some potential for the separation of different types of volume scattering based on their levels of depolarisation, they remain inferior to full [*S*] matrix systems.

4

Decomposition theorems

In Chapter 1 we saw how physical boundary conditions on a perfect conductor lead to a π phase shift on plane reflection, which when applied to the backscattering amplitude matrix leads to a canonical series of the form shown in equation (4.1):

$$[S]_{BSA} = \begin{bmatrix} 1 & 0\\ 0 & (-1)^{n+1} \end{bmatrix}$$
(4.1)

Here *n* is the order of reflection, so n = 1 is single reflection leading to the 2×2 identity matrix, while n = 2 is characteristic of dihedral reflections, with π phase difference between polarisations. These lead, via vectorization, to canonical Pauli scattering vectors of the form $(1,0,0)^{T}$ and $(0,1,0)^{T}$. These vectors form the building blocks for a generalization of this theory to encompass reflections from arbitrary dielectric interfaces, as we now demonstrate.

This generalization is of key practical importance for the development of applications in remote sensing, where variations in dielectric constant and angle of incidence must always be considered. In this chapter we see how to extend the vector formulation to deal with dielectric interfaces, with perfect conductors then arising as a special case. The general formulation then leads us to consider coherent decomposition theorems (Cloude, 1985; Krogager, 1993; Cameron, 1996). We develop these ideas in several stages, first by considering transformations—such as rotation of frame of coordinates or phase shifts to circular polarisations—and determine their impact on a description of reflection at dielectric boundaries. Secondly, we then generalize the vector formulation to cope with arbitrary boundary conditions in a transformation invariant way, which requires a generalization of the vector approach via the concept of the alpha parameter. Finally, we show how to incorporate depolarisation effects into this formulation by considering noncoherent decomposition theorems (Huynen, 1970; van Zyl, 1990; Cloude, 1996; Freeman, 1998; Yamaguchi, 2005).

4.1 Coherent decomposition theorems

We begin with a simple but important example. Consider how the scattering amplitude matrix [S] changes when we rotate the transmitter and receiver coordinate systems through an angle θ . In matrix form the rotation can be represented by a similarity transformation as shown in equation (4.2):

$$[S(\theta)] = \begin{bmatrix} \cos\theta & -\sin\theta\\ \sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} S_{HH} & S_{HV}\\ S_{VH} & S_{VV} \end{bmatrix} \begin{bmatrix} \cos\theta & \sin\theta\\ -\sin\theta & \cos\theta \end{bmatrix}$$
(4.2)

By expanding the matrices on either side in terms of the Pauli spin matrices we can write this as a unitary vector transformation as shown in equation (4.3) (verified by expansion and use of standard trigonometric identities):

$$\underline{k}(\theta) = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 0 & 0 & 0\\ 0 & \cos 2\theta & -\sin 2\theta & 0\\ 0 & \sin 2\theta & \cos 2\theta & 0\\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} S_{HH} + S_{VV}\\ S_{HH} - S_{VV}\\ S_{HV} + S_{VH}\\ S_{HV} - S_{VH} \end{bmatrix}$$
(4.3)

Here we note that the complex sums $S_{hh} + S_{vv}$ and $S_{hv} - S_{vh}$ are invariant to rotations, which already gives them a special physical significance. Of particular importance is the form of this transformation for backscatter problems, when the transmit and receive coordinates coincide and θ corresponds to rotation of the object about the line of sight. In this case, $S_{hv} = S_{vh}$ in the BSA system, and so for this case the transformation reduces to three dimensions, as shown in equation (4.4):

$$\underline{k}(\theta) = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 0 & 0\\ 0 & \cos 2\theta & -\sin 2\theta\\ 0 & \sin 2\theta & \cos 2\theta \end{bmatrix} \cdot \begin{bmatrix} S_{HH} + S_{VV}\\ S_{HH} - S_{VV}\\ 2S_{HV} \end{bmatrix} = [R]\underline{k}$$
(4.4)

Hence we see that while 4×4 unitary matrices $[U_4]$ are required to represent the most general bistatic scattering case, 3×3 unitary matrices $[U_3]$ are of importance for the special case of backscatter. In equation (4.3) we noted that $S_{\rm hh} + S_{\rm vv}$ is an invariant to rotation, and it is interesting to ask if there are any other complex combinations of the [S] matrix that remain unchanged by rotations. To see this we now consider the use of eigenvalue decompositions.

4.1.1 Roll invariance and eigenvectors

We first calculate the eigenvectors of the transformation matrix [R] in equation (4.4). By definition these states remain unchanged (apart from multiplication by a complex scalar, the eigenvalue) on operation of the matrix. Hence any linear combination of these states will also preserve its form. For the rotation matrix, calculation of the eigenvectors is straightforward, as shown in equation (4.5):

$$[R] \underline{u} = \lambda \underline{u} \Rightarrow \underline{u} = \begin{pmatrix} 1\\0\\0 \end{pmatrix}, \frac{1}{\sqrt{2}} \begin{pmatrix} 0\\1\\i \end{pmatrix}, \frac{1}{\sqrt{2}} \begin{pmatrix} 0\\1\\-i \end{pmatrix} \quad \lambda = 1, e^{-i2\theta}, e^{i2\theta} \quad (4.5)$$

Importantly, these three eigenvectors form a complete set; that is, any \underline{k} vector can be expanded uniquely in terms of this mutually orthogonal set. Hence we can use the matrices corresponding to these vectors as a basis for the expansion of an arbitrary [S] matrix to obtain equation (4.6):

$$\underline{k} = x \begin{pmatrix} 1\\0\\0 \end{pmatrix} + y \begin{pmatrix} 0\\1\\i \end{pmatrix} + z \begin{pmatrix} 0\\1\\-i \end{pmatrix} = f \begin{pmatrix} 1\\0\\0 \end{pmatrix} + g e^{-i2\theta} \begin{pmatrix} 0\\1\\i \end{pmatrix} + h e^{i2\theta} \begin{pmatrix} 0\\1\\-i \end{pmatrix}$$
(4.6)

where the complex scalars f, g and h remain invariant to rotations of the object about the line of sight. Furthermore, we may give each vector a physical interpretation in terms of a canonical scattering mechanism by converting the vectors back into their corresponding [S] matrices, as shown in equation (4.7).

$$[S] = f \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + ge^{-i2\theta} \begin{bmatrix} 1 & i \\ i & -1 \end{bmatrix} + he^{i2\theta} \begin{bmatrix} 1 & -i \\ -i & -1 \end{bmatrix}$$
$$= \begin{bmatrix} f + ge^{-i2\theta} + he^{i2\theta} & i(ge^{-i2\theta} - he^{i2\theta}) \\ i(ge^{-i2\theta} - he^{i2\theta}) & f - ge^{-i2\theta} - he^{i2\theta} \end{bmatrix}$$
(4.7)

The first of these corresponds to the sphere symmetric specular scattering process, while the remaining two correspond to scattering from a helix, the first (g) with right and the second (h) with left sense. Both scatter *all* incident waves into circular polarisation, and hence are not found generically as paired polarising objects in scattering by natural media. For this reason an alternative factorization is commonly used, designed to maintain invariance to rotations but employing more generic scattering mechanisms.

4.1.2 Krogager and Cameron decompositions

To motivate this idea we first rewrite the [S] matrix in the circular basis as shown in equation (4.8), and then extract the 'dominant' helix, leaving a remainder term as shown.

$$\begin{split} [S_{circ}] &= \begin{bmatrix} he^{i2\theta} & if \\ if & -ge^{-i2\theta} \end{bmatrix} \\ &= \begin{bmatrix} 0 & if \\ if & 0 \end{bmatrix} + \begin{cases} \begin{bmatrix} ge^{i2\theta} & 0 \\ 0 & -ge^{-i2\theta} \end{bmatrix} + \begin{bmatrix} (h-g)e^{i2\theta} & 0 \\ 0 & 0 \end{bmatrix} & \text{if } |h| > |g| \\ \begin{bmatrix} he^{i2\theta} & 0 \\ 0 & -he^{-i2\theta} \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & (g-h)e^{-i2\theta} \end{bmatrix} & \text{if } |g| > |h| \end{split}$$

$$(4.8)$$

The remainder term is the same in both cases, and has the form of a canonical scattering process: dihedral scattering at an angle θ . This expansion was first proposed in Krogager (1992, 1993), and has the vector form shown in equation (4.9). Here we see that while the sphere symmetric term has been maintained, the left and right helical coefficients have been combined into a single dominant helix scattering term and the remainder used to generate a new scattering mechanism: a dihedral scatterer oriented at angle $\theta/2$.

$$\underline{k} = f \begin{pmatrix} 1\\0\\0 \end{pmatrix} + \begin{cases} (h-g) \begin{pmatrix} 0\\e^{i2\theta}\\-ie^{i2\theta} \end{pmatrix} + 2g \begin{pmatrix} 0\\\cos 2\theta\\\sin 2\theta \end{pmatrix} \text{ if } |h| > |g| \\ (g-h) \begin{pmatrix} 0\\e^{-i2\theta}\\ie^{-i2\theta} \end{pmatrix} + 2h \begin{pmatrix} 0\\\cos 2\theta\\\sin 2\theta \end{pmatrix} \text{ if } |g| > |h| \end{cases}$$
(4.9)

This technique is often called the Krogager or SDH (sphere-diplane-helix) decomposition, in recognition of its use of these three basic building blocks to represent arbitrary scatterers. However, we note that these are still assumed to be three metallic scatterers, and we cannot yet account for interactions at dielectric interfaces. The three real non-negative components of this decomposition are directly obtained from the circular polarisation components, as shown in equation (4.10):

$$k_s = |S_{RL}| \quad k_D = \min(|S_{RR}| \quad |S_{LL}|) \quad k_H = \pm(|S_{RR}| - |S_{LL}|) \quad (4.10)$$

As well as providing the three real parameters of the SDH decomposition, this technique can also yield an efficient algorithm for estimation of the orientation of the scatterer θ (Lee, 2002; Schuler, 2002). From the form of the circular *S* matrix (equation (4.8)) we see that θ can be directly related to the product of S_{LL} and S_{RR} as shown in equation (4.11). In equation (4.11a) we show how the phase of the LL/RR product is related to θ and to the phase of hg^* .

$$S_{LL}S_{RR}^* = -hg^* e^{i4\theta}$$

$$\arg(S_{LL}S_{RR}^*) = \pi + \arg(hg^*) + 4\theta$$
(4.11a)

$$-S_{LL}S_{RR}^{*} = -(S_{HH} - S_{VV} + 2iS_{HV}) \left(S_{VV}^{*} - S_{HH}^{*} - 2iS_{HV}^{*}\right)$$

= $|S_{HH} - S_{VV}|^{2} - 4 |S_{HV}|^{2} + i4\text{Re}((S_{HH} - S_{VV}) S_{HV}^{*})$
 $\Rightarrow \tan^{-1}\left(\frac{4\text{Re}((S_{HH} - S_{VV}) S_{HV}^{*})}{|S_{HH} - S_{VV}|^{2} - 4 |S_{HV}|^{2}}\right) = \arg(hg^{*}) + 4\theta$ (4.11b)

$$\arg \langle -S_{LL}S_{RR}^* \rangle = \tan^{-1} \frac{4\operatorname{Re}\left(\langle (S_{HH} - S_{VV})S_{HV}^* \rangle\right)}{\langle |S_{HH} - S_{VV}|^2 \rangle - 4 \langle |S_{HV}|^2 \rangle} = 4\overline{\theta}$$
$$\Rightarrow \overline{\theta} = \frac{1}{4} \tan^{-1} \frac{4\operatorname{Re}\left(\langle (S_{HH} - S_{VV})S_{HV}^* \rangle\right)}{\langle |S_{HH} - S_{VV}|^2 \rangle - 4 \langle |S_{HV}|^2 \rangle} = \frac{1}{4} \tan^{-1} \frac{2\operatorname{Re}(t_{23})}{t_{22} - t_{33}}$$
(4.11c)

In equation (4.11b) we show how the phase of $-S_{LL}S_{RR}^*$ can then be related to the sum of 4θ and a scattering phase between *h* and *g*. Importantly, the product hg^* is often real, and so this phase is zero. We can see from equation (4.11b) that this will occur, for example, when S_{HV} equals zero. This always occurs for some θ for so-called symmetric scatterers, as discussed in the Cameron decomposition below. More generally, however, in the presence of depolarisation we can use averaging to estimate the average phase, as shown in equation (4.11c). Here we can see that the average scattering phase will be zero when we are aligned to the dominant axis in reflection symmetric depolarisers (see Section 2.4.2.3), since S_{HV} is then uncorrelated with both S_{HH} and S_{VV} . In this case we see from the equation that θ may then be easily related to the elements of the coherency matrix [*T*]. Note that while this is a robust and widely used algorithm for estimation of θ s its main drawback is the limited range of θ available, due to the factor of $\frac{1}{4}$ used in the equation.

This discussion about orientation leads us to consider an important class of objects termed symmetric scatterers (Cameron, 1996; Touzi, 2007). These



Fig. 4.1 Definition of symmetric scatterers

are defined as objects having an axis of symmetry in the plane of polarisation. When this axis is aligned with the antenna coordinates then there must be zero crosspolarisation, as shown schematically in Figure 4.1. For a horizontally polarised incident wave we see that secondary current induced above and below the symmetry line have parallel copolarised but antiparallel crosspolarised components. This leads to cancellation of the crosspolarised response of the object. Formally, the scattering matrix is formed as the coherent sum of two terms in the scattering matrix group $[S] + [S_{\beta}]$ in equation (1.145). The sum always results in a diagonal matrix in the HV system, as shown in Figure 4.1.

For such objects, therefore, there always exists a linear basis that diagonalizes the scattering amplitude matrix. In terms of the scattering vector it follows that it must be possible to find θ such that the crosspolar terms go to zero. This implies that the backscattering vector from such symmetric objects must be constrained, as shown in equation (4.12):

$$\begin{bmatrix} a'\\b'\\c' \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0\\0 & \cos 2\theta & -\sin 2\theta\\0 & \sin 2\theta & \cos 2\theta \end{bmatrix} \cdot \begin{bmatrix} a\\b\\c \end{bmatrix} \xrightarrow{c'=0} \underline{k} = \begin{bmatrix} \cos \alpha\\\sin \alpha \cos 2\theta e^{i\phi}\\\sin \alpha \sin 2\theta e^{i\phi} \end{bmatrix}$$
$$\Rightarrow \phi_{HH-VV} = \phi_{HV} \quad (4.12)$$

Here we see that it is only possible to diagonalize [S] by a rotation if the phase of the complex HV term equals that of the HH-VV term, which again gives special significance to the Pauli matrix expansion of [S]. This result provides us with an alternative algorithm for estimating the orientation angle of *symmetric* scatterers from the scattering vector, as shown in equation (4.13):

$$\tan 2\theta = \frac{2S_{HV}}{S_{HH} - S_{VV}} - \frac{\pi}{2} \le \theta \le \frac{\pi}{2}$$
(4.13)

While this appears to be a better approach to the LL/RR phase (it provides a wider range of angle estimates, for example) it is formed as the ratio of complex entities and hence is sensitive to fluctuations or departures from the symmetric assumption (and is unstable when $S_{\rm HH} = S_{\rm VV}$; that is, for specular surface scattering). However, there are two important ideas springing from this observation, both of which are concerned with devising strategies for dealing with scatterers that do not obey the symmetry constraint.

The first is called the Cameron decomposition (Cameron, 1996). The idea here is to filter the scattering vector *before* applying equation (4.13). The filtering is performed so as to generate the maximally symmetric component of the complex scattering vector—to keep as much of the original vector as possible, but to enforce the symmetry assumptions that the phase of HV and HH-VV are equal and that the amplitude ratio gives $\tan\theta$. The former requires that HV and HH-VV share a common complex factor ε , as shown in equation (4.14). To find ε we first calculate a residual vector $\Delta \underline{k}$ between the original and symmetric approximation, and then choose ε and θ to minimize the norm:

$$\Delta \underline{k} = \begin{bmatrix} 0\\ b - \varepsilon \cos 2\theta\\ c - \varepsilon \sin 2\theta \end{bmatrix} \xrightarrow{\theta, \varepsilon} \min \Delta \underline{k}^{*T} \Delta \underline{k}$$

$$\rightarrow \frac{\partial}{\partial \varepsilon} = 0 \Rightarrow \varepsilon - b \cos 2\theta - c \sin 2\theta = 0$$

$$\rightarrow \frac{\partial}{\partial \theta} = 0 \Rightarrow \left(|b|^2 - |c|^2 \right) \sin 4\theta - (bc^* + cb^*) \cos 4\theta = 0$$
(4.14)

The Cameron algorithm can then be presented in two stages, as shown in equation (4.15). The first is to generate an estimate of θ from the ratio of real quantities, and then to use this value to combine the complex numbers *b* and *c* into a single term ε .

$$\underline{k} = \begin{bmatrix} a \\ b \\ c \end{bmatrix} \xrightarrow{\text{max symmetic component}} \underline{k}_{sym} = \begin{bmatrix} a \\ \varepsilon \cos 2\theta \\ \varepsilon \sin 2\theta \end{bmatrix}$$
(4.15)
$$\tan 4\theta = \frac{bc^* + b^*c}{bb^* - cc^*} \quad \varepsilon = b\cos 2\theta + c\sin 2\theta$$

Note that the orientation estimate is then formally equivalent to the LL/RR coherence discussed in equation (4.11). One other important parameter deriving from this filtering idea is the degree of symmetry D_{sym} of a scattering vector, which expresses the ratio of power in the symmetric component to the total power, as shown in equation (4.16).

$$D_{sym} = \frac{|a|^2 + |\varepsilon|^2}{|a|^2 + |b|^2 + |c|^2} \quad 0 \le D_{sym} \le 1$$
(4.16)

The second approach to dealing with orientation of non-symmetric scattering vectors was first developed in Huynen (1987), and further developed in Touzi (2007). The idea is to rotate the vector until the scattering matrix can be expressed as the product of a diagonal form (in general the singular values of [S]) and an ellipticity or 'tau' transformation matrix (see equation (1.120)). This approach contains symmetric scatterers as a special case, when the ellipticity transformation becomes the identity matrix. The canonical (zero rotation) form of the [S] matrix in this approach is given by equation (4.17):

$$[S] = \begin{bmatrix} \cos \tau & i \sin \tau \\ i \sin \tau & \cos \tau \end{bmatrix} \cdot \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} \cdot \begin{bmatrix} \cos \tau & i \sin \tau \\ i \sin \tau & \cos \tau \end{bmatrix} \quad \lambda_1, \lambda_2 \in C \quad (4.17)$$

When we vectorize this expression using the Pauli basis we obtain the following canonical form:

$$\underline{k} = \begin{bmatrix} (\lambda_1 + \lambda_2)\cos 2\tau \\ (\lambda_1 - \lambda_2) \\ i(\lambda_1 + \lambda_2)\sin 2\tau \end{bmatrix} = \begin{bmatrix} \cos 2\tau & 0 & i\sin 2\tau \\ 0 & 1 & 0 \\ i\sin 2\tau & 0 & \cos 2\tau \end{bmatrix} \begin{bmatrix} \lambda_1 + \lambda_2 \\ \lambda_1 - \lambda_2 \\ 0 \end{bmatrix}$$
(4.18)

This can be used to estimate the unknown rotation of a general \underline{k} vector by recognising that the first and third elements of the canonical form are in phase quadrature so that the following relation gives an expression for θ :

$$\sin 2\theta \operatorname{Re}\left(\frac{b}{a}\right) + \cos 2\theta \operatorname{Re}\left(\frac{c}{a}\right) = 0 \Rightarrow \tan 2\theta = -\frac{\operatorname{Re}\left(\frac{c}{a}\right)}{\operatorname{Re}\left(\frac{b}{a}\right)}$$
(4.19)

While this is formally correct, again it does not allow extension to the idea of coherence and so is susceptible to complex noise fluctuations. For this reason the most robust algorithm for orientation estimation remains the LL/RR coherence (equation (4.11)). The only exception to this is in low noise (low entropy) environments or when the restricted range of the LL/RR estimate is too limiting for the application.

4.1.3 The scattering alpha parameter

We have seen above that the circular polarisation base provides a good platform from which to derive information about the mean orientation of the scatterer (from the LL/RR coherence phase). In Section 4.2.6 we show that circular polarisation is also a good base to assess the effects of mean Faraday rotation (LR/RL coherence phase). The scattering vector formulation in the Pauli base connects these different processes as special cases of 4×4 unitary transformations. This idea can be extended to consider the physical interpretation of general unitary transformations of the scattering vector. In particular we consider one important example: the alpha parameter, which will allow us finally to consider an approach to decomposition at dielectric boundaries. The problem with dielectrics is that they change the ratio of copolar [S] matrix elements. Hence we need to move away from a description based on the simple ratios implicit in the Pauli matrices; that is, we need now to consider not just rotations but unitary transformations that allow us to move smoothly away from the Pauli basis. This can be achieved by using transformations of the form shown in equation (4.20).

$$\underline{k}(\alpha) = \frac{1}{\sqrt{2}} \begin{bmatrix} \cos \alpha & -\sin \alpha & 0 & 0\\ \sin \alpha & \cos \alpha & 0 & 0\\ 0 & 0 & 1 & 0\\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} S_{HH} + S_{VV}\\ S_{HH} - S_{VV}\\ S_{HV} + S_{VH}\\ S_{HV} - S_{VH} \end{bmatrix}$$
(4.20)

We see that α changes the copolarised terms of the scattering matrix as required, although it remains invariant to rotations of the measurement coordinates. The angle is conveniently defined over the range $0 \le \alpha \le 90^\circ$, and represents a smooth change of scattering mechanism. To see this we start with the sphere symmetric specular mechanism with target vector $\underline{k} = [1,0,0,0]^T$ and corresponding 2×2 identity matrix for [S]. This we define as the $\alpha = 0^\circ$ boundary. Changes in α then correspond to departures in scattering from this reference mechanism. These variations can be summarized as shown in Figure 4.2. In the range $0^\circ \le \alpha \le 45^\circ$ the copolarised scattering terms differ in amplitude but are equal in phase, reaching, as an extreme point, $\alpha = 45^\circ$, which corresponds to an anisotropic scatterer with only one non-zero copolarised term. A simple example of this is a linear dipole scatterer, which has strong copolarised



Fig. 4.2 Physical interpretation of the scattering alpha angle

scatter for polarisations parallel to its axis and zero for the orthogonal case. In the range $45^{\circ} < \alpha \le 90^{\circ}$ we move into a region where the copolarised scattering coefficients are 180 degrees out of phase. This can occur for multiple scattering (dihedrals) or asymmetric scattering such as helices. In the extreme case of $\alpha = 90^{\circ}$ these scatterers have equal amplitude but maintain the 180-degree phase difference (metallic dihedrals or helices, for example). The alpha angle for any [S] matrix can then be directly estimated as shown in equation (4.21):

$$\cos \alpha = \frac{|S_{HH} + S_{VV}|}{\sqrt{2}\sqrt{|S_{HH}|^2 + 2|S_{HV}|^2 + |S_{VV}|^2}} \quad 0 \le \alpha \le \frac{\pi}{2}$$
(4.21)

One key advantage of using this angle in place of simple ratios such as $S_{\text{HH}}/S_{\text{VV}}$ is that it is invariant to rotations of the object. For example, a dipole changes its [S] matrix elements with rotation θ (see equation (1.161)), but both numerator and denominator of equation (4.21) remain invariant to θ , and hence α is unchanged. The same is true for rotation of a dihedral or any other rotationally dependent scatterer. Hence we can identify the scattering mechanism without needing to know the orientation. There remains an interesting question of how best to estimate the mean α angle in the presence of noise and fluctuations. This cannot be answered using a simple coherence estimate as it was for the orientation θ , and requires application of the entropy/alpha decompositions presented in Chapter 2.

To illustrate this we highlight an important special case of this alpha transformation. We can, for example, use it to represent the set of all diagonal [S] matrices with complex diagonal values λ_1 and λ_2 as SU(2) transformations of the identity (corresponding to a sphere symmetric scattering mechanism), as shown in equation (4.22):

$$\frac{1}{\sqrt{2}\sqrt{|\lambda_1|^2 + |\lambda_2|^2}} \begin{bmatrix} \lambda_1 + \lambda_2 \\ \lambda_1 - \lambda_2 \\ 0 \end{bmatrix} = e^{i\phi} \begin{bmatrix} \cos \alpha & -\sin \alpha e^{-i\delta} & 0 \\ \sin \alpha e^{i\delta} & \cos \alpha & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$
$$\Rightarrow \cos \alpha = \frac{|\lambda_1 + \lambda_2|}{\sqrt{2}\sqrt{|\lambda_1|^2 + |\lambda_2|^2}}, \delta = \arg\left(\frac{\lambda_1 - \lambda_2}{\lambda_1 + \lambda_2}\right) \tag{4.22}$$

We can now combine the above results to propose a 'point reduction theorem', as follows. According to this theorem we can express an arbitrary scattering mechanism \underline{w} (and here we restrict attention to the backscatter reciprocal threeelement case) as transformations of the sphere symmetric Pauli scatterer, as shown in equation (4.23):

$$\underline{w} = e^{i\phi} \begin{bmatrix} \cos \alpha \\ \sin \alpha \cos \psi e^{i\phi_1} \\ \sin \alpha \sin \psi e^{i\phi_2} \end{bmatrix}$$

$$= e^{i\phi} \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos 2\theta & -\sin 2\theta \\ \sin 2\theta & \cos 2\theta \end{bmatrix} \begin{bmatrix} \cos 2\tau & 0 & i \sin 2\tau \\ 0 & 1 & 0 \\ i \sin 2\tau & 0 & \cos 2\tau \end{bmatrix} \begin{bmatrix} \cos \alpha_d & -\sin \alpha_d e^{-i\delta} & 0 \\ \sin \alpha_d e^{i\delta} & \cos \alpha_d & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$
(4.23)

Starting on the right-hand side we begin with a 2×2 unitary sub-matrix that represents the change of (complex) singular values of the [S] matrix away from the identity matrix. We then invoke a transformation to account for ellipticity in the singular polarisation states—a departure from symmetric scatterers. Finally we include a rotation of the scatterer in the plane of polarisation. On the left-hand side we show the standard trigonometric parameterization of unitary vectors in three dimensions. We note the following relationships:

1. The parameter α remains invariant to rotations as expected, but actually satisfies the following relation to parameters of the point reduction transform:

C

$$\cos \alpha = \cos 2\tau \cos \alpha_d \tag{4.24}$$

Only in the case of symmetric scatterers ($\tau = 0$) does $\alpha = \alpha_d$. Nonetheless, α remains a useful parameter as a measure of the departure of the scattering process from sphere symmetry.

2. The parameter ψ is only equal to the rotation angle in case of symmetric scatterers ($\tau = 0$). In other cases the ψ parameter contains a mixture of dependence on rotation and ellipticity. In the non-symmetric case, orientation is better estimated using equation (4.19).

4.1.4 Orthogonal scattering mechanisms

Very often we wish to determine the component of a scattering vector \underline{k} in a predetermined 'direction' \underline{w} . To do this we can form the projection of \underline{k} onto \underline{w} as the Hermitian inner product, as shown in equation (4.25):

$$s = \underline{w}^{*T}\underline{k} = re^{i\phi} \tag{4.25}$$

The power in this projection can then be computed as shown in equation (4.26):

$$p = ss^* = \underline{w}^{*T} \underline{k} \underline{k}^{*T} \underline{w} = r^2 \tag{4.26}$$

It is also of interest to consider the match between \underline{w} and the scattering mechanism embedded in \underline{k} . This can be found from the normalized inner product, as shown in equation (4.27):

$$d = \frac{\underline{w}^{*T}\underline{k}}{|\underline{k}|} \quad 0 \le |d| \le 1 \tag{4.27}$$

Clearly, if the scattering mechanisms are the same then d = 1. Interestingly, however, we can also obtain 'orthogonality' between mechanisms, defined when d = 0. In this case the power in the projection will be zero. Orthogonality is already a powerful concept in the analysis of elliptically polarised waves, which allowed us to construct bases from paired combinations of orthogonal polarisations. Through vectorization of [S] we can now extend this analysis to scattering mechanisms themselves.

It is interesting to investigate the orthogonality of the canonical mechanisms identified in the various decompositions in the previous section. For example, the Pauli matrix set form an orthogonal set, as their vectors are all mutually orthogonal. Hence, in addition to obvious orthogonalities such as horizontal and vertical dipoles, we can also count spheres and 45° and 0° rotated dihedrals as orthogonal mechanisms. The eigenvectors of the rotation matrix [*R*] used in equation (4.5) also form an orthogonal set. However, the SDH components themselves do not form an orthogonal set (see equation (4.9)). The helix and dihedral do not represent orthogonal processes.

This orthogonality is a desirable property of coherent decomposition theorem (CDT) schemes, as it provides optimum separability in classification schemes and forms the basis for an expansion of second-order averages in the coherency matrix, as discussed in Section 4.2. To illustrate this we consider a sphere mapping based on diagonal [S] matrices, as developed in Chapter 3. In this case we consider classification of scattering vectors of the reduced form, shown in equation (4.28):

$$\underline{w}_{1} = \begin{bmatrix} \cos \alpha \\ \sin \alpha e^{i\delta} \\ 0 \end{bmatrix} \Rightarrow \quad \underline{w}_{\perp} = \begin{bmatrix} \sin \alpha \\ -\cos \alpha e^{i\delta} \\ 0 \end{bmatrix}$$
(4.28)

This is mathematically identical to the case of orthogonal waves in C2 and consequently we can map each scattering mechanism as a point on the scattering sphere (analogous to the Poincaré sphere) and define for each mechanism an orthogonal one, as shown in equation (4.28). These two mechanisms will have antipodal points on the sphere, and α and δ represent the spherical triangle coordinates of the points. This result has two important consequences. The first is that we can define each point on the sphere using a real three-vector—the equivalent of the Stokes vector for waves. The second that is we can define a coherency matrix between the two complex singular values and extract a mean scattering mechanism in the presence of noise. We now turn to consider the machinery required to deal with fluctuations in the [S] matrix elements, and consider the general problem of depolarisation. First, however, we consider the issue of orthogonality at dielectric interfaces.

4.1.5 Orthogonality of scattering mechanisms in natural terrain

A key question arising from the concept of scattering matrix orthogonality is its probable occurrence in scattering from natural media. This is particularly important in remote sensing applications when we often wish to separate depolarising surface and volume scattering components. In particular, we saw in Chapter 3 that either one of the dihedral or direct surface scattering mechanisms can be dominant, or indeed both can occur simultaneously in scattering from natural terrain, depending on the topography, vegetation structure and surface roughness. In this section we look at the degree to which these processes remain orthogonal over a wide range of variations in dielectric constant.

If we define \underline{w}_B as the scattering vector for Bragg surface backscattering at angle θ from a surface with dielectric constant ε_r , and likewise the corresponding dihedral mechanisms (at the same angle and dielectric constant of the surface as \underline{w}_D), then the degree of orthogonality Q can be defined (in dB) from the magnitude of an inner product, as shown in equation equation (4.29):

$$Q = 20 \log_{10} \left(\left| \underline{w}_D^{*T} \underline{w}_B \right| \right) \tag{4.29}$$



Fig. 4.3 Schematic of combined direct and specular dihedral surface returns

The smaller Q, the better the isolation between the mechanisms. Small Q better enables separation of the mechanisms when both are present in the same scattering problem, as they then emerge as separate eigenvectors of the coherency matrix [T]. The scenario we are considering is an arbitrary combination of surface and dihedral scattering, as shown schematically in Figure 4.3. Note that the relative amplitude of the mechanisms is not so important in this context as is their relative polarimetric properties. Figures 4.4 and 4.5 show the factor Q for the same two angles of incidence used in Figures 3.15 and 3.16. Note that Q is smaller for the larger angle of incidence. For the shallower angle, Qcan still be small (-10 dB or so), but for dry materials (with a small $\varepsilon_r < 20$)





Fig. 4.4 Magnitude of Q (equation (4.23)) for 22.5-degree angle of incidence

Fig. 4.5 Magnitude of Q (equation (4.23)) for 45-degree angle of incidence

at shallow angles we note that the separation is the poorest at around $-6 \, \text{dB}$. Again we note, for the shallow angle, that it is the dielectric constant of the second surface B that causes the largest variation in Q.

The larger angle of incidence has a wide range of combinations of dielectric constants where Q is better than -20 dB. Hence we note that separability of the two mechanisms \underline{w}_D and \underline{w}_B , based on orthogonality, is largely independent of dielectric constant variations and improves with increasing angle of incidence.

This discussion of orthogonality, of course, has the largest impact in depolarisation problems, when the orthogonal mechanisms can be easily separated as eigenvectors of the Hermitian coherency or covariance matrix. This brings us to consider the set of issues raised by incoherent decompositions.

4.2 Incoherent decomposition theorems

We have seen that it is often useful to represent the scattering coherency matrix [T] (or equivalently the Mueller matrix [M]) as a sum of composite elements, so in general we can write an expansion of the form shown in equation (4.30):

$$[T] = \sum_{i=1}^{R} [T_i] \Leftrightarrow [M] = \sum_{i=1}^{R} [M_i]$$
(4.30)

As a simple example, consider a signal + noise decomposition as shown in equation (4.31), where n is the noise variance.

As a point of caution we note again that *n* may be caused by depolarisation (by a Lambertian surface, for example) as well as by system noise effects. Only in rank-3 scattering (reciprocal backscattering, for example) can we estimate the system noise *n* from the difference between crosspolarisation channels as the smallest eigenvalue of the HV/VH 2×2 coherency matrix shown in equation (4.32) (Hajnsek, 2001):

$$\hat{n} = \frac{1}{2} \left(\left\langle S_{HV} S_{HV}^* \right\rangle + \left\langle S_{VH} S_{VH}^* \right\rangle - \left\langle S_{VH} S_{VH}^* \right\rangle \right)^2 + 4 \left\langle S_{HV} S_{VH}^* \right\rangle \left\langle S_{VH} S_{HV}^* \right\rangle \right)$$
(4.32)

This can then be used as a method of noise filtering, for example, by inverting equation (4.31):

$$[T_1] = [T] - \hat{n} [T_N] \Leftrightarrow [M_1] = [M] - \hat{n} [M_N]$$
(4.33)

In this section we look at generalizations and extensions of this idea. Such representations are called incoherent decomposition (ICDs), because the addition of coherency matrices implies addition of power, which in turn implies there is no phase coherence between the elements. When this is not the case we must employ coherent decomposition theorems (CDTs)—a description of which was given in the previous section. There are two key elements to any ICD:

- 1. The choice of order R in equation (4.30). In principle we are free to choose any order, but in practice there are several constraining factors. The first is recognition of the fact that each composite matrix must have at least one free parameter to describe its form and structure (n for the noise element in equation (4.31), for example). On the other hand we have seen that a general coherency or Mueller matrix has at most sixteen free parameters, and so this means that R < 16. Further restrictions apply from various symmetry constraints. The most important of these is the reciprocity theorem in backscatter, for which the coherency matrix has rank 3 and only nine free parameters. In most backscatter applications, therefore, $R \leq 9$ (a notable exception being the treatment of backscatter Faraday rotation, which is nonreciprocal, and has $R \leq 10$). Other symmetries may apply on a case-by-case basis: for example, backscatter reflection symmetry (a common case in microwave remote sensing), for which R < 6, and the most severe case, being backscatter azimuth symmetry with R < 2.
- 2. The nature of the component elements of the expansion T_i . This leads to a further general partitioning of the decomposition as follows. The idea is often to force one or more components of the ICD to correspond to a rank-1 coherency matrix (or equivalently a Mueller matrix that corresponds to a single amplitude [S] matrix). Consequently, we can write the general ICD in the form shown in equation (4.34), where $P \le R$ is the number of rank-1 components:

$$[T] = \sum_{i=1}^{P} [T_i] + \sum_{j=Q}^{R} [T_j] \Leftrightarrow [M] = \sum_{i=1}^{P} [M_i] + \sum_{j=Q}^{R} [M_j] \quad (4.34)$$

Each rank-1 component requires up to a maximum of seven $(4 \times 4$ coherency) or five $(3 \times 3$ coherency) parameters. If each component actually requires Q_i parameters, then it follows that $Q = \sum_{i=1}^{P} Q_i + 1$.

A natural extension of this concept is to consider each element of the expansion a reflection symmetric depolariser. This then represents the most general case found in the literature. There are two important special cases of this approach, as follows.

4.2.1 The Huynen decomposition: rank 1 scattering + noise decompositions

It is interesting to speculate on the existence of a generalization of the signal + noise decomposition in equation (4.31). In particular we now seek decomposition into a single rank 1 component (the 'signal') $[T_1]$ plus a remainder $[T_N]$

(considered as 'noise'), as shown in equation (4.35):

$$[T] = [T_1] + [T_N] \Leftrightarrow [M] = [M_1] + [M_N]$$
(4.35)

There certainly exist interesting special cases of this type of decomposition, as first pointed out by Chandrasekhar (1960) in his work on radiative transfer theory. Following this example we consider lateral scattering (scattering through 90° bistatic angle) by a random cloud of small spheroids (equation (3.101)) to obtain a coherency matrix of the form shown in equation (4.36), where A_p is the ratio of polarizabilities or particle shape parameter. As shown, this can then be reformulated as a rank-1 scattering mechanism (dipole scattering <u>k</u>_D) plus a 'noise' term, although in this case the 'noise' is not thermal but is caused by depolarisation due to the particle anisotropy (for example, the noise term goes to zero for spherical particles, $A_p = 1$).

$$\begin{split} [T] &= \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 & 0\\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 & 0\\ 0 & 0 & 1 & 0\\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \lambda_1 & 0 & 0 & 0\\ 0 & \lambda_2 & 0 & 0\\ 0 & 0 & \lambda_2 & 0\\ 0 & 0 & 0 & \lambda_2 \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & 0 & 0\\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 & 0\\ 0 & 0 & 1 & 0\\ 0 & 0 & 0 & 1 \end{bmatrix} \\ &= (\lambda_1 - \lambda_2) \underline{k}_D \underline{k}_D^{*T} + \lambda_2 \begin{bmatrix} 1 & 0 & 0 & 0\\ 0 & 1 & 0 & 0\\ 0 & 0 & 1 & 0\\ 0 & 0 & 0 & 1 \end{bmatrix} \\ &= \frac{1}{15} (2A_P^2 + 6A_P + 7) \underline{k}_D \underline{k}_D^{*T} + \frac{1}{15} (A_P - 1)^2 \begin{bmatrix} 1 & 0 & 0 & 0\\ 0 & 1 & 0 & 0\\ 0 & 0 & 0 & 1 \end{bmatrix}$$
(4.36)

Problems arise, however, when we try to generalize this idea to an arbitrary coherency matrix. The basic issue is that there is an infinite number of ways of extracting a rank 1 scattering mechanism from [T]. To see this, consider starting with an arbitrary mechanism \underline{w} . We can then always write the decomposition into a single mechanism plus remainder, as shown in equation (4.37):

$$[T] = t_w \underline{ww}^{*T} + [T_N] \quad t_w = \underline{w}^{*T} [T] \underline{w}$$
(4.37)

Some authors (most notably in the Huynen decomposition (Huynen, 1970, 1987)) have proposed solving this problem by insisting that the remainder term satisfy some invariance properties under a change of polarisation base. This then constrains the choice of \underline{w} for the single mechanism. To see how this works we begin by noting the importance of the null-space of the remainder matrix. By definition, $[T_N]$ lies in a subspace of the full polarisation space. Hence there always exist vectors \underline{v} , spanning the null-space of the matrix, defined as shown in equation (4.38):

$$[T_N] \underline{\nu} = 0 \tag{4.38}$$

Now, insistence that the remainder matrix $[T_N]$ maintains its form under a unitary change of base (that its null space is unchanged) requires that the null space \underline{v} is constrained to be an eigenvector of the change of basis matrix, as shown in equation (4.39):

$$[U_B][T_N][U_B]^{*T} \underline{v} = 0 \Rightarrow [U_B] \underline{v} = \lambda \underline{v}$$
(4.39)

Huynen, for example, considered reciprocal backscatter problems (a rank-3 coherency matrix) and insisted that the remainder was invariant to rotations about the line of sight through an angle θ ; that is, he assumed an explicit form for the change of base matrix, as shown in equation (4.40):

$$[U_B] = \begin{bmatrix} 1 & 0 & 0\\ 0 & \cos\theta & -\sin\theta\\ 0 & \sin\theta & \cos\theta \end{bmatrix} \Rightarrow \underline{\nu} = \begin{bmatrix} 1\\ 0\\ 0 \end{bmatrix}, \frac{1}{\sqrt{2}} \begin{bmatrix} 0\\ 1\\ i \end{bmatrix}, \frac{1}{\sqrt{2}} \begin{bmatrix} 0\\ 1\\ -i \end{bmatrix}$$
(4.40)

It follows that there are still three choices for the null space, given by the three eigenvectors of the change of basis matrix. Huynen selected the first on physical grounds and suggested the following general form, which now bears his name: the Huynen decomposition:

$$[T] = [T_1] + [T_N] = \begin{bmatrix} t_{11} & t_{12} & t_{13} \\ t_{12}^* & t_{22} - n_{22} & t_{23} - n_{23} \\ t_{13}^* & t_{23}^* - n_{23}^* & t_{33} - n_{33} \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 \\ 0 & n_{22} & n_{23} \\ 0 & n_{23}^* & n_{33} \end{bmatrix}$$
(4.41)

We clearly see the null space in the remainder term corresponding to the first eigenvector in equation (4.40), and the elements n_{ij} are selected so as to force the first matrix to have rank 1. However, by choosing one of the other eigenvectors in equation (4.40) for the null space we would produce a *different* rank 1 matrix in the expansion, and hence we see that although this approach does constrain selection of the rank-1 component it does not provide a unique solution. This approach does, however, raise the issue of orthogonal spaces and their importance in ICDs. We now turn to consider an alternative formulation, which solves the uniqueness problem by considering a general eigenvalue expansion.

4.2.2 The Cloude–Pottier decomposition

The discussion in the previous section pointed out several uniqueness problems in attempts to extract a single 'dominant' rank-1 coherency matrix from a general depolarising system. Paradoxically it is easier to expand any depolarising system as a sum of four rank-1 components, as shown in equation (4.42):

$$[T] = t_1 [T_1] + t_2 [T_2] + t_3 [T_3] + t_4 [T_4]$$

$$\Leftrightarrow [M] = m_1 [M_1] + m_2 [M_2] + m_3 [M_3] + m_4 [M_4]$$
(4.42)

Cloude and Pottier first proposed this idea in Cloude (1996), and the decomposition therefore bears their name. Proof of the uniqueness of this expansion follows directly from the Hermitian nature of the coherency matrix, which can be expanded in terms of its orthogonal eigenvectors and real eigenvalues, as shown in equation (4.43):

$$[T] = \sum_{i=1}^{4} \lambda_i \underline{e}_i \underline{e}_i^{*T} \begin{cases} \lambda_1 \ge \lambda_2 \ge \lambda_3 \ge \lambda_4 \ge 0 \in \mathfrak{R} \\ \underline{e}_i^{*T} \underline{e}_j = 0 \end{cases}$$
(4.43)

There are four important ideas stemming from this approach.

4.2.2.1 Eigenvector decomposition: dominant scattering mechanisms

We see that the decomposition in equation (4.42) is unique if we associate $t_i = \lambda_i [T_i] = \underline{e}_i \underline{e}_i^{*T}$. It follows that if we want to identify a 'dominant' and unique rank-1 scattering mechanism we should choose the eigenvector corresponding to the maximum eigenvalue of the coherency matrix (Cloude, 1989). In this way we can finally write a basis invariant and unique expansion into a rank-1 plus noise decomposition, as shown in equation (4.44):

$$[T] = \lambda_1 [T_1] + [T_N] \quad [T_N] = \sum_{i=2}^{4} \lambda_i \underline{e}_i \underline{e}_i^{*T}$$
(4.44)

In this case the null space of $[T_N]$ is defined by the dominant eigenvector \underline{e}_1 , and by definition this remains invariant to *all* unitary transformations of the problem, and not just the rotation invariance required in the Huynen approach. We note that this approach also has the advantage of automatically scaling with varying rank of the coherency matrix. In backscatter reciprocity, for example, the smallest eigenvalue is always zero, and the most general expansion is therefore provided in terms of three components, as shown in equation (4.45). The extension to rank 2 and rank 1 depolarising systems follows immediately.

$$[T] = t_1 [T_1] + t_2 [T_2] + t_3 [T_3] \Leftrightarrow [M] = m_1 [M_1] + m_2 [M_2] + m_3 [M_3]$$
(4.45)

4.2.2.2 Eigenvector secomposition and contrast optimization

This eigenvector expansion has several useful applications other than identifying the dominant rank-1 scattering process. For example, a common problem in radar imaging is the enhancement of contrast between two scatterers (Novak, 1989). In this case we consider two depolarising systems with corresponding coherency matrices $[T_A]$ and $[T_B]$, and wish to find the single scattering mechanism <u>w</u> that maximizes the contrast between the two systems. The contrast can be defined as a ratio of intensities, given by the ratio of Hermitian forms Q, as shown in equation (4.46). Optimization of contrast can then be formulated using the Lagrangian shown, and the single scattering mechanism that maximizes this ratio is obtained as an eigenvector of the product of matrices, as shown in equation (4.46):

$$Q = \frac{\underline{w}^{*T} [T_A] \underline{w}}{\underline{w}^{*T} [T_B] \underline{w}} \xrightarrow{\max} L = \underline{w}^{*T} [T_A] \underline{w} - \lambda(\underline{w}^{*T} [T_B] \underline{w} - 1)$$

$$\Rightarrow \frac{\partial L}{\partial \underline{w}} = [T_A] \underline{w} - \lambda [T_B] \underline{w} = 0 \Rightarrow [T_B]^{-1} [T_A] \underline{w} = \lambda \underline{w}$$
(4.46)

In this case, therefore, the best single mechanism to isolate in each matrix is \underline{w} , and not the individual eigenvectors \underline{e}_i . In this case we make use of the following ICD:

$$[T_A] = t_A \underline{w} \underline{w}^{*T} + [T_{NA}] \quad t_A = \underline{w}^{*T} [T_A] \underline{w}$$

$$[T_B] = t_B \underline{w} \underline{w}^{*T} + [T_{NB}] \quad t_B = \underline{w}^{*T} [T_B] \underline{w}$$

(4.47)

4.2.2.3 Eigenvector decomposition and CFAR detection

As a second important example of eigenvector decompositions we consider the problem of radar backscatter detection in the presence of depolarising 'clutter'; that is, unwanted background signal (Ioannidis, 1979; Novak, 1989; Wanielik, 1992). In this case we make a measurement of the scattering matrix [S], and wish to set a threshold in order that the false alarm rate be constant (that is the rate of error associated with making a detection when actually no target is present) in the presence of depolarising Gaussian clutter (see Appendix 3), forming a so-called constant-false-alarm-rate, or CFAR, detector. In this case we first vectorize a sample into the form \underline{k} as shown in equation (4.48), and then assume for the clutter background alone that the vector has a multivariate normal distribution as shown. In this case the CFAR implementation involves calculating a metric Q as shown, and deciding on detection with constant errors based on the setting of a threshold 'x'. Note that the metric Q involves the inverse coherency matrix for the background depolarising system, and so assumes we have some knowledge of this matrix, either by measurement or modelling.

$$[S] \xrightarrow{\text{vectorise}} \underline{k} = \begin{bmatrix} a \\ b \\ c \end{bmatrix} \Rightarrow p(\underline{k}) = \frac{1}{\pi^3 \det([T])} e^{-k^T [T]^{-1} \underline{k}}$$
$$\Rightarrow Q = \underline{k}^{*T} [T]^{-1} \underline{k} > x \qquad (4.48)$$

It is instructive to formulate this detection process in terms of the eigenvector decomposition of the coherency matrix. Equation (4.49) shows how Q can be expressed as a function of the eigenvalues and eigenvectors of [T].

$$Q = |\underline{k}|^{2} \underline{w}^{*T} \left(\frac{1}{\lambda_{1}} \underline{e}_{1} \underline{e}_{1}^{*T} + \frac{1}{\lambda_{2}} \underline{e}_{2} \underline{e}_{2}^{*T} + \frac{1}{\lambda_{3}} \underline{e}_{3} \underline{e}_{3}^{*T} \right) \underline{w}$$

$$= \frac{|\underline{k}|^{2}}{\lambda_{1}} \left| \underline{w}^{*T} . \underline{e}_{1} \right|^{2} + \frac{|\underline{k}|^{2}}{\lambda_{2}} \left| \underline{w}^{*T} . \underline{e}_{2} \right|^{2} + \frac{|\underline{k}|^{2}}{\lambda_{3}} \left| \underline{w}^{*T} . \underline{e}_{3} \right|^{2}$$
(4.49)

We now see an interesting case where the optimum solution lies in the smallest eigenvalue. The best detection arises when \underline{k} (with corresponding mechanism \underline{w}) is orthogonal to the first and second eigenvalues ($\underline{w}^{*T}\underline{e}_i = 0$) and lies parallel to the smallest eigenvector $\underline{w}^{*T}\underline{e}_3 \approx 1$. In this case (λ_3 small) the contrast between the 'target' and 'clutter' is maximized, and the Q factor is large even when the absolute value of \underline{k} is small, leading to so-called sub-clutter visibility of the target. In this case it is best to characterize the depolarising 'clutter' using the following ICD:

$$[T] = \lambda_{\min} \underline{e}_3 \underline{e}_3^{*T} + [T_N] \tag{4.50}$$

4.2.2.4 The entropy/alpha decomposition

The three previous applications have all sought to identify a single eigenvalue/ eigenvector to solve an optimization problem. However, we have seen in Section 2.4 that we can also identify an *average* scattering mechanism from the eigenvectors and associate it with a degree of disorder, the entropy. This gives us an invariant two-parameter characterization of the depolariser represented by [T]. Furthermore, we have seen that we can then plot such paired values in a plane: the entropy/alpha diagram (see Section 2.4.2.4). This can then be used to classify different types of scattering behaviour. This approach has been applied, for example, as an unsupervised classification method in imaging radar polarimetry (Pottier, 1997; Cloude, 1997a), and as a physics based pre-processor for more general statistical classification techniques based, for example, on the Wishart distribution (Ferro-Famil, 2001; Lee, 1999, 2008).

These examples all illustrate the importance of considering the full eigenvalue spectrum of the coherency matrix, and how in some applications there is useful information even in the smaller eigenvalues and eigenvectors. Another way to systematically exploit this information is to consider an approach to ICDs based on physical scattering models, as we now consider.

4.2.3 Model-based incoherent decompositions

One important alternative approach to ICDs is to use physical models of scattering depolarisation to determine the number and parameterization of each component (van Zyl, 1989; Freeman, 1998; Dong, 1998; Yamaguchi, 2005). The general starting point for this approach is to identify the main components to any backscattering scenario, as summarized schematically in Figure 4.6. This model in its simplest form involves a generic two-layer approach, with a volume layer of particles (vegetation, snow, and so on) above a non-penetrable boundary or surface. The total average backscatter is then determined from a 3×3 coherency matrix composed of three main elements, as shown in equation (4.51). The first component—S in Figure 4.6—is direct backscatter from the underlying surface, recognising that its scattering behaviour is modified by propagation through the top layer, which may cause attenuation of the surface response but can also act to distort the polarisation behaviour of the surface. For relatively smooth surfaces the small perturbation or Bragg scattering model is often used to model this component. (Although alternatives such as the X-Bragg or IEM can be easily included (see equation 3.42), they always incur the expense of adding more parameters and hence make inversion more difficult.)

$$[T] = [T_S] + [T_D] + [T_V]$$

$$= m_{S} \begin{bmatrix} \cos^{2} \alpha_{s} & \sin \alpha_{s} \cos \alpha_{s} e^{i\phi_{s}} & 0\\ \sin \alpha_{s} \cos \alpha_{s} e^{-i\phi_{s}} & \sin^{2} \alpha_{s} & 0\\ 0 & 0 & 0 \end{bmatrix} + m_{D} \begin{bmatrix} \cos^{2} \alpha_{d} & \sin^{2} \alpha_{d} & 0\\ \sin \alpha_{d} \cos \alpha_{d} e^{-i\phi_{d}} & \sin^{2} \alpha_{d} & 0\\ 0 & 0 & 0 \end{bmatrix} + m_{V} \begin{bmatrix} 2A_{P}^{2} + 6A_{P} + 7 & 0 & 0\\ 0 & (A_{P} - 1)^{2} & -i4\kappa(A_{P} - 1)^{2}\\ 0 & i4\kappa(A_{P} - 1)^{2} & (A_{P} - 1)^{2} \end{bmatrix}$$
(4.51)

As shown in equation (3.36), Bragg scattering acts as a strong polariser with zero depolarisation and (for a flat surface) zero crosspolarisation. Thus Bragg scattering contributes a rank-1 coherency matrix with three unknown parameters, as shown in equation (4.51), where $\alpha_s < \pi/4$ depends on dielectric constant and angle of incidence. The second component—D in Figure 4.6—involves a multiple scattering interaction between the surface reflection and



Fig. 4.6 Schematic representation of the various scattering contributions used in modelbased decompositions



Fig. 4.7 Geometry of dihedral backscattering mechanism

volume-scattering elements that return the reflected signal back to the observer. While this combination can often be complex and a source of wave depolarisation, it is often assumed dominated by a double specular reflection process or simple dihedral reflection from dielectrics, rather than a reflection/scattering combination. For this to occur the volume needs to be populated not only by small particles (leaves, branches, and so on) but also by some electrically large scatterers (vertical tree-trunks, for example), in which case the reflection from the surface (modelled by the Fresnel equations, modified by the Rayleigh coefficient to account for surface roughness; see equation (3.31)) is followed by a second Fresnel reflection from the surface of the large scatterer, as shown schematically in Figure 4.7. This reflection can be written in terms of the angle of incidence using the constraint that the final angle of reflection must be 180° for backscatter, shown in equation (4.52):

$$[S_D] \propto \begin{bmatrix} R_{H1}R_{H2} & 0\\ 0 & -R_{V1}R_{V2} \end{bmatrix} \rightarrow \begin{cases} R_{H1} = \frac{\cos\theta - \sqrt{\varepsilon_{r1} - \sin^2\theta}}{\cos\theta + \sqrt{\varepsilon_{r1} - \sin^2\theta}}\\ R_{V1} = \frac{\varepsilon_{r1}\cos\theta - \sqrt{\varepsilon_{r1} - \sin^2\theta}}{\varepsilon_{r1}\cos\theta + \sqrt{\varepsilon_{r1} - \sin^2\theta}}\\ R_{H2} = \frac{\sin\theta - \sqrt{\varepsilon_{r2} - \cos^2\theta}}{\sin\theta + \sqrt{\varepsilon_{r2} - \cos^2\theta}}\\ R_{V2} = \frac{\varepsilon_{r2}\sin\theta - \sqrt{\varepsilon_{r2} - \cos^2\theta}}{\varepsilon_{r2}\sin\theta + \sqrt{\varepsilon_{r2} - \cos^2\theta}} \end{cases}$$
(4.52)

In this case the dihedral component can be modelled as a rank-1 polarising element with three parameters, as shown in equation (4.51), with $\alpha_d > \pi/4$ depending now on the angle of incidence and the two dielectric constants of surface and reflector.

The third component—V in Figure 4.6—is direct volume scattering from the top layer itself. By assuming this layer has azimuthal symmetry it can be modelled as a random cloud of spheroidal particles with shape parameter A_p , and chirality κ , the coherency matrix for which was derived in equation (3.97). This contribution acts as a three-parameter depolariser with coherency matrix, as shown in equation 4.51. An obvious extension is to include multiple scattering as discussed in Section 3.4.3.

In total, this model generates (at least) nine parameters for only six observations or measurements (the three diagonal elements of the coherency matrix plus one off-diagonal complex and one imaginary term). This is important because we often try to use such models for inversion or parameter estimation; that is, to compare the predictions of the model against the measurement and adjust the parameters until the difference between the two is minimized. We then take the parameters used for the minimization as estimates of the true values. However, when the number of unknowns exceeds the number of observations we face problems of uniqueness; that is, many different combinations of the parameters could achieve a match against the limited measurements, and so we are unable to select the true solution. It is therefore of interest, in such model-based ICDs, to try to more closely match the number of model parameters to the number of measurements. There are, of course, many ways to do this, but several common assumptions emerge in practice.

4.2.4 The Freeman–Durden decomposition

One common set of approximations appears in the Freeman–Durdan decomposition, as follows (Freeman, 1998). The first idea is to ignore volume chirality ($\kappa = 0$), and also to select a fixed value for A_p —the particle shape in the volume term. This is often assumed to correspond to a cloud of prolate spheroids or dipoles ($A_p \gg 1$). This then reduces the unknowns to seven. The second feature is then to assume that in practice one or other of the surface or dihedral responses dominates, and so we can set the minor mechanism to have a known α value ($\alpha_s = 0$ if the dihedral is dominant, or $\alpha_s = \pi/2$ if the surface is dominant) without too much loss of accuracy. This again reduces the number of unknown model parameters by two. In this way we obtain a balanced system with five model parameters and five unknowns, which can now be inverted. In summary, the Freeman–Durden decomposition—a special case of equation (4.51), and an example of an ICD into two rank-1 components plus one rank-3—can be written in terms of the coherency matrix, as shown in equation (4.53):

$$[T] = \begin{bmatrix} t_{11} & t_{12} & 0 \\ t_{12}^* & t_{22} & 0 \\ 0 & 0 & t_{33} \end{bmatrix}$$
$$= m_v \begin{bmatrix} 2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} + m_s \begin{bmatrix} \cos^2 \alpha_s & \cos \alpha_s \sin \alpha_s e^{i\phi_s} & 0 \\ \cos \alpha_s \sin \alpha_s e^{-i\phi_s} & \sin^2 \alpha_s & 0 \\ 0 & 0 & 0 \end{bmatrix}$$
$$+ m_d \begin{bmatrix} \cos^2 \alpha_d & \cos \alpha_d \sin \alpha_d e^{i\phi_d} & 0 \\ \cos \alpha_d \sin \alpha_d e^{-i\phi_d} & \sin^2 \alpha_d & 0 \\ 0 & 0 & 0 \end{bmatrix} \xrightarrow{\text{either}} \begin{cases} \alpha_s = 0, \phi_s = 0 \\ \alpha_d = 0, \phi_d = 0 \end{cases}$$
$$(4.53)$$

As mentioned, the main reason for adopting this simplified form is to enable inversion of the model directly from data, as follows. The first step is to use the HV channel to directly estimate the volume scattering component m_v , as shown in equation 4.54:

$$m_v = t_{33}$$
 (4.54)

We can then estimate the α parameter, under the hypothesis of a dominant scattering mechanism, by first calculating its apparent value from the coherency matrix (here called the Freeman alpha value, α_F) and then assigning the model parameters to α_F according to a $\pi/4$ threshold, as shown in equation (4.55):

$$\tan \alpha_{F} = \frac{t_{22} - t_{33}}{\sqrt{t_{12} \cdot t_{12}^{*}}} \Rightarrow \begin{cases} \alpha_{d} = \alpha_{F}, \alpha_{s} = 0 \\ \Rightarrow m_{d} = \frac{t_{22} - t_{33}}{\sin^{2} \alpha_{d}}, m_{s} = t_{11} - 2t_{33} - m_{d} \cos^{2} \alpha_{d} & \text{if } \alpha_{F} > \frac{\pi}{4} \\ \alpha_{s} = \alpha_{F}, \alpha_{d} = \frac{\pi}{2} \\ \Rightarrow m_{s} = \frac{t_{11} - 2t_{33}}{\cos^{2} \alpha_{s}}, m_{d} = t_{22} - t_{33} - m_{d} \sin^{2} \alpha_{s} & \text{if } \alpha_{F} < \frac{\pi}{4} \end{cases}$$
(4.55)

Also shown in equation (4.55) are equations used to estimate the scattering power terms m_s and m_d . Finally, the last parameter, the scattering phase, can be estimated as shown in equation (4.56):

$$\begin{cases} \phi_d = \arg(t_{12}), \phi_s = 0 & \text{if } \alpha_F > \frac{\pi}{4} \\ \phi_s = \arg(t_{12}), \phi_d = 0 & \text{if } \alpha_F < \frac{\pi}{4} \end{cases}$$
(4.56)

One attractive feature of this model is that it represents a direct decomposition of the total backscattered power into three components, as shown in equation (4.57):

$$P_{total} = P_S + P_D + P_V = m_s + m_d + 4m_v \tag{4.57}$$

Note, however, that the model inversion does not guarantee non-negative estimates of power. In the presence of noise or violations of the underlying model assumptions, some of the power components could be estimated as negative—an undesirable feature of such models.

4.2.5 Generalized Freeman–Durden decompositions

The Freeman–Durden approach is clearly not unique, and we can choose to simplify the general model of equation (4.51) in many different ways. Several examples exist in the literature: for example, Freeman modified his own approach (Freeman 2007) by relaxing the random assumption for the volume scattering and instead assigning the particles an orientation distribution with a single parameter to be estimated from the data. A second approach is that due to Yamaguchi (2005). The main additive features of his technique are to include non-symmetric features such as scattering from helical type scatterers, as occurs in chiral materials, or to consider more general volume scattering terms, allowing for a relaxation of the azimuthal symmetry assumption to reflection symmetry. Rather than present an exhaustive survey of these methods, we return to the key idea of target orthogonality to show how the eigenvector approach can be used to enhance such models.

In the Freeman approach nothing was mentioned about orthogonality of the surface and dihedral component, but we have seen that for any surface mechanism with parameter α_s the orthogonal scattering mechanism has $\alpha = \pi/2 - \alpha_s$, which lies in the dihedral regime. Therefore, another approach to reducing the unknowns in equation (4.51) is to postulate that the surface and dihedral mechanisms are orthogonal. This allows us to express α_d in terms of α_s and thus reduce the parameter count by one. When added to the relaxed Freeman assumption of azimuthal symmetric scattering for the volume term and the loss of one phase angle, we again obtain a balanced system with five parameters and five measurements. This model—which we can call a hybrid Freeman/eigenvalue technique—has the compact form shown in equation (4.58), where α is now determined as predominantly surface or dihedral scattering, depending on the dominant eigenvalue of a 2 × 2 sub-matrix as shown, where $F_P = 2$ corresponds

to the Freeman-Durden model.

$$\begin{aligned} [T] &= m_s \begin{bmatrix} \cos \alpha \\ -\sin \alpha \\ 0 \end{bmatrix} \cdot \begin{bmatrix} \cos \alpha & -\sin \alpha & 0 \end{bmatrix} + m_d \begin{bmatrix} \sin \alpha \\ \cos \alpha \\ 0 \end{bmatrix} \cdot \begin{bmatrix} \sin \alpha & \cos \alpha & 0 \end{bmatrix} + m_v \begin{bmatrix} F_p & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \\ &= \begin{bmatrix} m_s \cos^2 \alpha + m_d \sin^2 \alpha & \cos \alpha \sin \alpha (m_d - m_s) & 0 \\ \cos \alpha \sin \alpha (m_d - m_s) & m_d \cos^2 \alpha + m_s \sin^2 \alpha & 0 \\ 0 & 0 & 0 \end{bmatrix} + \begin{bmatrix} F_p m_v & 0 & 0 \\ 0 & m_v & 0 \\ 0 & 0 & m_v \end{bmatrix} \\ &\Rightarrow [T]_{\text{SD}} = [T] - m_v \begin{bmatrix} F_p & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \\ &= \begin{bmatrix} \cos \alpha & \sin \alpha & 0 \\ -\sin \alpha & \cos \alpha & 0 \\ 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} m_s & 0 & 0 \\ 0 & m_d & 0 \\ 0 & 0 & 0 \end{bmatrix} \cdot \begin{bmatrix} \cos \alpha & -\sin \alpha & 0 \\ \sin \alpha & \cos \alpha & 0 \\ 0 & 0 & 1 \end{bmatrix} . \end{aligned}$$
(4.58)

This model can then be inverted by first calculating m_s and m_d as eigenvalues of the rank 2 matrix T_{SD} , and then using these to obtain estimates of the α parameter by first parameterising the eigenvector as $(1,e_2)^T$, then solving for the complex number e_2 , and finally normalizing the eigenvector by $\sqrt{1 + |e_2|^2}$ to obtain α , as shown in equation (4.59):

$$m_{v} = t_{33}$$

$$m_{d,s} = \frac{(t_{11} + t_{22} - (F_{p} + 1)t_{33}) \pm \sqrt{(t_{11} - t_{22} - (F_{p} - 1)t_{33})^{2} + 4|t_{12}|^{2}}}{2}$$

$$\alpha_{d,s} = \cos^{-1} \left[\left(1 + \left| \frac{t_{12}}{t_{22} - t_{33} - m_{d,s}} \right|^{2} \right)^{-\frac{1}{2}} \right]$$
(4.59)

Note also that this approach provides a means for avoiding negative powers by keeping only the non-negative eigenvalue spectrum of T_{SD} . Also, if we assume one of m_s , m_d is always zero then we can estimate the volume parameter F_p . We can also use this model to estimate the ratio of surface-to-volume scattering components, termed μ , which, as we shall see in Chapter 7, is an important parameter in the development of polarimetric interferometry. According to the model, when we select the crosspolarisation channel HV, the surface-to-volume scattering ratio μ is by definition zero, as the surface components are all constrained into the rank-2 upper portion of the coherency matrix. On the other hand, to find the surface-to-volume ratio for the largest surface component we can proceed as follows. First we select the maximum 'surface' component from the maximum of the pair m_s and m_d as defined in equation (4.59). (Note that we are including both direct surface and dihedral returns in the 'surface' components to distinguish them from the pure volume scattering component.) From this we can then calculate the alpha parameter of the maximum component, as shown in equation (4.60):

$$m_{\max} = \max(m_d, m_s) \Rightarrow \alpha_{\max} = \cos^{-1} \left[\left(\sqrt{1 + \left| \frac{t_{12}}{t_{22} - t_{33}} - m_{\max} \right|^2} \right)^{-\frac{1}{2}} \right]$$
(4.60)

If we then select a scattering mechanism \underline{w} based on this alpha parameter, we can obtain an expression for the desired ratio, as shown in equation (4.61):

$$[T] = m_{\max} \begin{bmatrix} \cos \alpha_{\max} \\ -\sin \alpha_{\max} \\ 0 \end{bmatrix} \cdot [\cos \alpha_{\max} & -\sin \alpha_{\max} & 0] \\ + m_{\min} \begin{bmatrix} \sin \alpha_{\max} \\ \cos \alpha_{\max} \\ 0 \end{bmatrix} \cdot [\sin \alpha_{\max} & \cos \alpha_{\max} & 0] \\ + m_{\nu} \begin{bmatrix} F_p & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \\ \underline{w} = \begin{bmatrix} \cos \alpha_{\max} \\ \sin \alpha_{\max} \\ 0 \end{bmatrix} \Rightarrow \underline{w}^{*T} \langle [T] \rangle \underline{w} = m_{\max} + m_{\nu} (1 + (F_p - 1) \cos^2 \alpha_{\max}) \\ \Rightarrow \mu = \frac{surface}{volume} = \frac{m_{\max}}{m_{\nu}} \frac{1}{(1 + (F_p - 1) \cos^2 \alpha_{\max})}$$
(4.61)

Notice that this ratio is not the *maximum* surface-to-volume ratio. This can always be obtained by using a <u>w</u> calculated as a generalized eigenvector of the matrix $[T_V]^{-1}[T_{SD}]$ (see equation (4.46)). The general solution to this is complicated, but if we make the assumption that there is one dominant surface component (either m_s or m_d), then we can obtain a direct estimate of the maximum surface-to-volume ratio by finding the largest eigenvalue of $[T_V]^{-1}[T_{max}]$, as shown in equation (4.62):

$$\frac{m_{\max}}{m_{\nu}} \left| \frac{\cos^{2} \alpha_{\max}}{F_{p}} - \mu \frac{\cos \alpha_{\max} \sin \alpha_{\max}}{F_{p}} \right| = 0$$

$$\Rightarrow \mu_{\max} = \frac{m_{\max}}{m_{\nu}} \left(\sin^{2} \alpha_{\max} + \frac{1}{F_{p}} \cos^{2} \alpha_{\max} \right)$$
(4.62)

It is also interesting to note that we can also define another small value of surface-to-volume scattering (in addition to the HV channel) by selecting the \underline{w} vector orthogonal to the maximum surface component. In this case we obtain a μ value, given as shown in equation (4.63):

$$\underline{w} = \begin{bmatrix} -\sin \alpha_{\max} \\ \cos \alpha_{\max} \\ 0 \end{bmatrix} \Rightarrow \underline{w}^{*T} [T] \underline{w} = m_{\min} + m_{\nu} (1 + (F_p - 1) \sin^2 \alpha_{\max})$$

$$\Rightarrow \mu_{\min} = \frac{surface}{volume} = \frac{m_{\min}}{m_{\nu} (1 + (F_p - 1) \sin^2 \alpha_{\max})}$$
(4.63)

In the introduction to the general model of equation (4.51) we mentioned that the surface response is actually modified by propagation through the upper medium. So far, however, we have ignored such propagation effects, on the assumption that the volume is random and hence acts as a scalar attenuation of all polarisations equally, which has no effect on the alpha parameters of surface and dihedral components. However, in some cases this assumption is not true, and leads us to yet another class of decompositions—but this time based on a multiplicative expansion. We now turn to consider such ideas.

4.2.6 Propagation distortions in model-based decompositions

To consider the effects of wave propagation on depolarising systems (Azzam, 1978), we begin by considering the simplest case of an homogeneous propagation channel (see Section 1.2.4), in which the orthogonal eigenpolarisations for propagation in the medium are aligned with the [S] matrix measurement basis a,b. In this case there is by definition no cross-coupling between eigenstates, and so the propagation matrix is diagonal. The effect of wave propagation into the medium followed by scattering, and then propagation out of the medium, can then be represented as a matrix product, as shown in equation (4.64):

$$[S]_{observed} = \begin{bmatrix} e^{-i\beta_a z} & 0\\ 0 & e^{-i\beta_b z} \end{bmatrix} \cdot \begin{bmatrix} S_{aa} & S_{ab}\\ S_{ba} & S_{bb} \end{bmatrix} \cdot \begin{bmatrix} e^{-i\beta_a z} & 0\\ 0 & e^{-i\beta_b z} \end{bmatrix}$$

$$\begin{cases} \beta_a = \beta_o n_a = \beta_o n'_a - i\kappa_a\\ \beta_b = \beta_o n_b = \beta_o n'_b - i\kappa_b \end{cases}$$
(4.64)

where $\beta_o = \frac{2\pi}{\lambda}$ is the free space wavenumber. Note that β_a and β_b are in general complex, relating to both phase shifts and attenuation by the medium. It is convenient, in what follows, to factor the propagation matrix by extracting the mean propagation constant, as shown in equation (4.65):

$$\begin{bmatrix} \exp(-i\beta_a z) & 0\\ 0 & \exp(-i\beta_b z) \end{bmatrix}$$
$$= \exp\left(\frac{-i(\beta_a + \beta_b)}{2}z\right) \begin{bmatrix} \exp\left(-i\frac{(\beta_a - \beta_b)}{2}z\right) & 0\\ 0 & \exp\left(i\frac{(\beta_a - \beta_b)}{2}z\right) \end{bmatrix}$$
(4.65)

The first step in understanding the effects of propagation on scattering entropy is then to vectorize this matrix expression. The best way to do this is in the lexicographic basis, when we obtain the form shown in equation (4.66):

$$\underline{k}_{L}(z) = e^{-i(\beta_{a} + \beta_{b})z} \begin{bmatrix} \exp(vz) & 0 & 0 & 0\\ 0 & 1 & 0 & 0\\ 0 & 0 & 1 & 0\\ 0 & 0 & 0 & \exp(-vz) \end{bmatrix} \begin{bmatrix} S_{aa} \\ S_{ab} \\ S_{ba} \\ S_{bb} \end{bmatrix} = [P_{L}] \cdot \underline{k}_{L}$$

$$v = -i(\beta_{a} - \beta_{b}) \tag{4.66}$$

Note that the crosspolarised channels are influenced by the mean propagation constant, while the copolar channels are also influenced by the differential propagation. We can combine these effects into a single 4×4 matrix $[P_L]$ as

shown. The final step is then to express the matrix $[P_L]$ in the Pauli base, so we can finally relate propagation to the coherency matrix formulation. This is shown in equation (4.67), where we see we can express the matrix in terms of elementary hyperbolic functions.

$$[P_P] = \frac{1}{2} \begin{bmatrix} 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & -1 \\ 0 & 1 & 1 & 0 \\ 0 & i & -i & 0 \end{bmatrix} [P_L] \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & -i \\ 1 & -1 & 0 & 0 \end{bmatrix}$$
$$= e^{-i(\beta_a + \beta_b)z} \begin{bmatrix} \cosh \nu z & \sinh \nu z & 0 & 0 \\ \sinh \nu z & \cosh \nu z & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$
(4.67)

As an important special case, if the medium is random (if it has azimuthal symmetry), then the differential propagation constant is zero, and the propagation matrix $[P_P]$ reduces to a multiple of the identity matrix. This is the form implicit in the Freeman decomposition. In this case the surface component m_S is attenuated by the mean extinction in the medium, but its polarimetry remains unchanged. A second important example is when the medium is lossless (zero extinction) but acts as a retarder (with differential phase shifts). In this case ν is purely imaginary, and the propagation matrix has the special form shown in equation (4.68):

$$[P_{\rm P}] = e^{-i(\beta_a + \beta_b)z} \begin{bmatrix} \cosh ivz & \sinh ivz & 0 & 0\\ \sinh ivz & \cosh ivz & 0 & 0\\ 0 & 0 & 1 & 0\\ 0 & 0 & 0 & 1 \end{bmatrix}$$
$$= e^{-i(\beta_a + \beta_b)z} \begin{bmatrix} \cos(vz) & i\sin(vz) & 0 & 0\\ i\sin(vz) & \cos(vz) & 0 & 0\\ 0 & 0 & 1 & 0\\ 0 & 0 & 0 & 1 \end{bmatrix}$$
(4.68)

In general, however, the coherency matrix of the scatterers beneath the propagation channel will have a distorted form, as shown in equation (4.69):

$$[T(z)] = [P_P][T][P_P]^{*T} = e^{-2(\kappa_b + \kappa_a)z} [P(\tau)][T][P(\tau^*)]$$

$$\tau = \upsilon z = ((\kappa_b - \kappa_a) - i\beta_0(n'_a - n'_b))\frac{z'}{\cos\theta_0}$$
(4.69)

where z is now the slant range coordinate, related to z', the vertical coordinate, by θ_0 , the angle of incidence (see Figure 3.38). We note again the attenuation of amplitude by mean wave extinction, but also note that since P is not generally diagonal then there results some distortion to the apparent polarimetric parameters of [*T*].

To illustrate this we consider the effects on the apparent scattering parameter α for a scatterer viewed through a medium with zero retardence but total



Fig. 4.8 Distortion of the alpha parameter due to differential extinction in the volume layer

differential extinction Δ . By changing variable we can then express (from equation (4.67)) the propagation transformation directly in terms of the differential extinction in dB and in terms of the tangent of the new alpha parameter, as shown in equation (4.70) (see equation 3.11):

$$\begin{bmatrix} \cos \alpha' \\ \sin \alpha' \end{bmatrix} = e^{-s} \begin{bmatrix} \cosh A & \sinh A \\ \sinh A & \cosh A \end{bmatrix} \cdot \begin{bmatrix} \cos \alpha \\ \sin \alpha \end{bmatrix}$$

$$\Rightarrow \tan \alpha' = \frac{\sin \alpha + \cos \alpha \cdot \tanh(0.1151 * \Delta)}{\cos \alpha + \sin \alpha \cdot \tanh(0.1151 * \Delta)}$$
(4.70)

Figure 4.8 shows how this distortion behaves for varying differential extinction. We see that for $\Delta = 0 \, dB$ we have zero distortion as expected (this is the basic assumption of the Freeman decomposition). However, as Δ increases we see that the range of apparent alpha becomes compressed until, when the differential extinction becomes very large, we obtain an apparent alpha of 45° for all scatterers. This corresponds to all scattering being filtered by a polariser with an orientation given by the dominant eigenpolarisation.

Finally, we consider the more general case when the eigenpolarisations are no longer matched to the backscattering matrix basis. In this case we must modify equation (4.63) to include a unitary congruent transformation by a 2×2 unitary matrix, as shown in equation (4.71):

$$\begin{bmatrix} S \end{bmatrix}_{observed} = \begin{bmatrix} e^{-i\beta_a z} & 0 \\ 0 & e^{-i\beta_b z} \end{bmatrix} \begin{bmatrix} U_2 \end{bmatrix} \cdot \begin{bmatrix} S_{aa} & S_{ab} \\ S_{ba} & S_{bb} \end{bmatrix} \begin{bmatrix} U_2 \end{bmatrix}^T \cdot \begin{bmatrix} e^{-i\beta_a z} & 0 \\ 0 & e^{-i\beta_b z} \end{bmatrix}$$
$$\begin{cases} \beta_a = \beta_o n'_a - i\kappa_a \\ \beta_b = \beta_o n'_b - i\kappa_b \end{cases}$$
(4.71)
This leads to a modified vectorization of the problem, as shown in equation (4.72):

$$\underline{k}_{L}(z) = e^{-i(\beta_{a} + \beta_{b})z} [P_{L}] \begin{bmatrix} U_{4}^{B} \end{bmatrix} \begin{bmatrix} S_{aa} \\ S_{ab} \\ S_{ba} \\ S_{bb} \end{bmatrix} = [P_{L}^{B}] \underline{k}_{L}$$
(4.72)

where the 4 × 4 unitary matrix $[U_4^B]$ is a function of two parameters: α_w and δ , defining the coordinates of the eigenpolarisation on the Poincaré sphere. It can be obtained by direct expansion, as shown in equation (4.73):

$$\begin{bmatrix} U_2 \end{bmatrix} = \begin{bmatrix} \cos \alpha_w & -\sin \alpha_w e^{-i\delta} \\ \sin \alpha_w e^{i\delta} & \cos \alpha_w \end{bmatrix} \Rightarrow \begin{bmatrix} U_2 \end{bmatrix} \otimes \begin{bmatrix} U_2 \end{bmatrix} = \begin{bmatrix} U_4^L \end{bmatrix}$$
$$\begin{bmatrix} U_4^L \end{bmatrix} = \begin{bmatrix} \cos^2 \alpha_w & -\cos \alpha_w \sin \alpha_w e^{-i\delta} & -\cos \alpha_w \sin \alpha_w e^{-i\delta} \\ \cos \alpha_w \sin \alpha_w e^{i\delta} & \cos^2 \alpha_w & -\sin^2 \alpha_w & -\cos \alpha_w \sin \alpha_w e^{-i\delta} \\ \cos \alpha_w \sin \alpha_w e^{i\delta} & -\sin^2 \alpha_w & \cos^2 \alpha_w & -\cos \alpha_w \sin \alpha_w e^{-i\delta} \\ \sin^2 \alpha_w e^{i2\delta} & \cos \alpha_w \sin \alpha_w e^{i\delta} & \cos \alpha_w \sin \alpha_w e^{i\delta} & \cos^2 \alpha_w \end{bmatrix}$$
$$(4.73)$$

Now converting to the Pauli basis we obtain the following form for the unitary matrix, as shown in equation (4.74):

$$\begin{bmatrix} U_4^P \end{bmatrix} = \frac{1}{2} \begin{bmatrix} 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & -1 \\ 0 & 1 & 1 & 0 \\ 0 & i & -i & 0 \end{bmatrix} \begin{bmatrix} U_4^L \end{bmatrix} \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & -i \\ 1 & -1 & 0 & 0 \end{bmatrix}$$
$$= \begin{bmatrix} \cos^2 \alpha_w + \sin^2 \alpha_w \cos 2\delta & i \sin^2 \alpha_w \sin 2\delta & 2i \sin \alpha_w \cos \alpha_w \sin \delta & 0 \\ -i \sin^2 \alpha_w \sin 2\delta & \cos^2 \alpha_w - \sin^2 \alpha_w \cos 2\delta & -2 \sin \alpha_w \cos \alpha_w \cos \delta & 0 \\ 2i \sin \alpha_w \cos \alpha_w \sin \delta & 2 \sin \alpha_w \cos \alpha_w \cos \delta & \cos^2 \alpha_w - \sin^2 \alpha_w & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$
(4.74)

Note, for example, when $\delta = 0$ —when we consider only linear eigenpolarisations—then this matrix reduces to the standard rotation matrix, as shown in equation (4.75):

$$\begin{bmatrix} U_4^P \end{bmatrix} = \begin{bmatrix} \cos^2 \alpha_w + \sin^2 \alpha_w & 0 & 0 & 0 \\ 0 & \cos^2 \alpha_w - \sin^2 \alpha_w & -2\sin \alpha_w \cos \alpha_w & 0 \\ 0 & 2\sin \alpha_w \cos \alpha_w & \cos^2 \alpha_w - \sin^2 \alpha_w & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$
$$= \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos 2\alpha_w & -\sin 2\alpha_w & 0 \\ 0 & \sin 2\alpha_w & \cos 2\alpha_w & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$
(4.75)

We can represent the distorting effects of wave propagation through an arbitrary (homogeneous) layer on the backscattering coherency matrix, as shown in equation (4.76):

$$[T(z)] = e^{-(\kappa_b + \kappa_a)z} [P(\tau)] [U_4^P(\alpha_w, \delta)] [T] [U_4^P(\alpha_w, \delta)]^{*T} [P(\tau^*)]$$

$$\tau = \nu z = \left((\kappa_b - \kappa_a) - i\beta_o \left(n'_a - n'_b \right) \right) \frac{z'}{\cos \theta_o}$$
(4.76)

Finally we turn to consider an important practical example of the utility of the vectorization scheme: to analyse the effects of Faraday propagation rotation distortion on target decomposition. Such effects must be considered, for example, in the analysis of low-frequency radar Earth observation data collected from space when propagation through the ionosphere cannot be ignored (see Section 1.2.2.3).

4.2.6.1 Scattering vector formulation of Faraday rotation

The scattering vector formulation can also be used to further analyse the effects of Faraday rotation (see equation (1.58)) (Bickel, 1965; Wright, 2003; Freeman, 2004). In this case the scattering matrix is transformed by the *same* rotation matrix on the left and right sides, as shown in equation (4.77):

$$\begin{bmatrix} S_{\psi} \end{bmatrix} = \begin{bmatrix} \cos\psi & \sin\psi \\ -\sin\psi & \cos\psi \end{bmatrix} \begin{bmatrix} S_{HH} & S_{HV} \\ S_{VH} & S_{VV} \end{bmatrix} \begin{bmatrix} \cos\psi & \sin\psi \\ -\sin\psi & \cos\psi \end{bmatrix}$$
(4.77)

By expansion and vectorization in terms of the Pauli matrices this can be written as a unitary transformation of the scattering vector, as shown in equation (4.78):

$$\underline{k}(\psi) = \frac{1}{\sqrt{2}} \begin{bmatrix} \cos 2\psi & 0 & 0 & -\sin 2\psi \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ \sin 2\psi & 0 & 0 & \cos 2\psi \end{bmatrix} \begin{bmatrix} S_{HH} + S_{VV} \\ S_{HH} - S_{VV} \\ S_{HV} + S_{VH} \\ S_{HV} - S_{VH} \end{bmatrix}$$
(4.78)

This relationship can be used to estimate the Faraday rotation by recognising that for calibrated reciprocal backscattering, $S_{\rm HV}$ – $S_{\rm VH}$ is zero on the right-hand side of this equation. Hence any observed difference between $S_{\rm HV}$ and $S_{\rm VH}$ must be due to Faraday effects, and from equation (4.78) it follows that the Faraday rotation can be found from the ratio shown in equation (4.79):

$$\tan 2\psi = \frac{S_{HV} - S_{VH}}{S_{HH} + S_{VV}}$$
(4.79)

One problem with practical use of equation (4.79), however, is that it involves the ratio of two complex quantities on the right and a real quantity on the left, and hence it is susceptible to data fluctuations and residual calibration errors. It would be more robust to devise an algorithm that involves the ratio of real quantities so that averaging may be used. We can use the scattering vector formulation combined with a 4×4 unitary transformation into the circular basis to solve this as follows. We begin by relating the lexicographic circular basis scattering vector to the Pauli expansion by a 4×4 unitary matrix, as shown in equation (4.80):

$$\begin{bmatrix} S_{LL} \\ S_{LR} \\ S_{RL} \\ S_{RR} \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} 0 & 1 & i & 0 \\ i & 0 & 0 & 1 \\ i & 0 & 0 & -1 \\ 0 & -1 & i & 0 \end{bmatrix} \frac{1}{\sqrt{2}} \begin{bmatrix} S_{HH} + S_{VV} \\ S_{HH} - S_{VV} \\ S_{HV} + S_{VH} \\ S_{HV} - S_{VH} \end{bmatrix} = [U_{4circ}] \underline{k} \quad (4.80)$$

This can be derived from the circular change of base transformation in equation (1.162). This can then be used to transform equation (4.80) into the circular basis by a unitary similarity transformation, as shown in equation (4.81):

$$\begin{bmatrix} S_{LL}(\psi) \\ S_{LR}(\psi) \\ S_{RL}(\psi) \\ S_{RR}(\psi) \end{bmatrix} = \begin{bmatrix} U_{4circ} \end{bmatrix} \begin{bmatrix} \cos 2\psi & 0 & 0 & -\sin 2\psi \\ 0 & 1 & 0 & 0 \\ \sin 2\psi & 0 & 0 & \cos 2\psi \end{bmatrix} \begin{bmatrix} U_{4circ} \end{bmatrix}^{*T} \begin{bmatrix} S_{LL} \\ S_{LR} \\ S_{RL} \\ S_{RR} \end{bmatrix}$$
$$= \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \exp(-i2\psi) & 0 & 0 \\ 0 & 0 & \exp(i2\psi) & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} S_{LL} \\ S_{LR} \\ S_{RL} \\ S_{RL} \\ S_{RL} \\ S_{RR} \end{bmatrix}$$
(4.81)

Note, significantly, that the copolarised terms (LL and RR) are not distorted by Faraday rotation. If again we assume that the underlying scatterer is reciprocal ($S_{LR} = S_{RL}$ on the right-hand side of equation (4.81)), as in backscatter BSA problems, then the Faraday rotation angle can simply be found from the phase of the product of crosspolarised terms, as shown in equation (4.82):

$$\psi = \frac{1}{4} \arg(S_{RL} S_{LR}^*) \tag{4.82}$$

which lends itself to averaging in the presence of data fluctuations by estimating the complex coherence between the cross-channels LR and RL (see Chapter 2 for a general discussion of coherence in polarimetry). Therefore, if the Faraday distorted linear scattering matrix $[S^F]$ is measured by a sensor, then the average Faraday rotation component may be obtained directly from equation (4.82), as shown explicitly in equation (4.83):

$$\tan 4\overline{\psi} = \left\langle \frac{-2\text{Re}((S_{HV}^F - S_{VH}^F)(S_{HH}^F + S_{VV}^F)^*)}{\left|S_{HH}^F + S_{VV}^F\right|^2 - \left|S_{HV}^F - S_{VH}^F\right|^2} \right\rangle$$
(4.83)

Note the mathematical similarity to the estimation of orientation using copolarised circular channels (equation (4.11)), although the two formulae describe very different physical phenomena.

In general we have seen that circular polarisation provides the best basis for a description of orientation effects (both reciprocal and non-reciprocal). Polarimetric phase and coherence play an important role in such studies. However, we have seen that in scattering from random media, entropy effects act to destroy such phase information and hence reduce our ability to exploit coherence. There is, however, a second way to generate phase information from active systems: by the use of interferometry. Here we shall see that new possibilities arise for the control of phase, even in random media problems, and this, when combined with polarimetry, will lead us finally to a general formulation of how best to exploit the 'orientation memory' effect in scattering from random media.

5

Introduction to radar interferometry

We saw in Chapters 3 and 4 that the polarisation properties of scattering from natural media, especially when volume scattering is involved, can lead to significant levels of depolarisation (high scattering entropy). This depolarisation has two characteristic features. Firstly, it can be very high (entropies above 0.9, even for single scattering), which limits the accuracy of polarimetric phase estimation due to the high speckle associated with high entropy (see Appendix 3 and López-Martínez (2005)). The second and more important issue is that the level of depolarisation is determined entirely by the structure of the medium and is independent of the sensor. There is no sensor control of entropy. On the other hand, radar interferometry, as we shall show, provides a means of controlling entropy by choice of sensor configuration (baseline). This, when combined with polarimetry, then gives us the ability to tune a sensor for a particular environment and to obtain good phase estimates, even in the presence of strong depolarisation. Before considering the combination of polarimetry with interferometry, we first review the basic features of single-channel radar interferometry (Bamler, 1998; Franceschetti, 1999; Kampes, 2006).

5.1 Radar interferometry

We have seen that a scattered single-channel radar signal may be represented by a complex scalar, representing its amplitude a and phase ϕ , as shown in equation (5.1):

$$s_1 = a_1 e^{i\phi_1} = a_1 e^{-i(\beta 2R_1 + \phi_{S1})} = a_1 e^{-i\left(\frac{2\pi}{\lambda} 2R_1 + \phi_{S1}\right)}$$
(5.1)

Here we have further decomposed the phase of the signal into two parts—the first a propagation phase that depends on the distance between the radar and the scattering point (R_1), and the second a scattering phase that depends on the detailed nature of the scattering process. In polarimetry we concentrated on looking at the changes of scattering phase by combining two signals with different polarisations but collected at the same point in space, so cancelling the range phase term, as shown in equation (5.2):

$$\left. \begin{array}{l} s_1 = a_1 e^{-i\left(\frac{4\pi}{\lambda}R_1 + \phi_{51}\right)} \\ s_2 = a_2 e^{-i\left(\frac{4\pi}{\lambda}R_1 + \phi_{52}\right)} \end{array} \right\} \Rightarrow s_1 s_2^* = a_1 a_2 e^{i\left(\phi_{52} - \phi_{51}\right)} \tag{5.2}$$

We have seen in the previous chapters how to characterize fluctuations in the polarimetric phase due to depolarisation in the scattering process. In particular we have seen that in the presence of random volume scattering from anisotropic particles this depolarisation can be very high (especially in the case of a random volume of extremely prolate spheroids), leading to a noisy polarimetric phase. It would be useful to be able to counter this depolarisation in some way. Radar interferometry provides such an option, as we now show.

In radar interferometry we begin again with the reference signal s_1 of equation (5.1). However, rather than diversify polarisation for a fixed position, we now keep polarisation constant but vary spatial position to collect a second signal s_2 from a displaced point, separated from the first by a spatial baseline B, as shown in Figure 5.1. Note that there are inherently two ways to achieve this. The first-'single-pass interferometry'-occurs when we use two radars at the same instant in time to collect signals s_1 and s_2 . Alternatively, it is often more convenient (and more cost effective) to use a single radar system to collect a signal at position 1 and then move the radar to position 2 at some time later to collect s_2 . In the airborne/space radar context this is called 'repeat-pass interferometry', and the time difference between the data collections is called the temporal baseline T_B (in seconds), to be contrasted with the spatial baseline B (with units of metres). Note that we have another degree of freedom for 'single pass' sensors in choosing either to transmit and receive from both points (the dual transmitter mode described in Figure 5.1 and often termed 'ping-pong'), or just to transmit from 1 and receive on both 1 and 2 (called the single transmitter or standard mode). In this case we can avoid the expense and complexity of a second transmitter. However, in this case the factor of 4π in equation (5.3) is halved to 2π , since the phase difference now only depends on $R_1 - R_2$ rather than $2R_1 - 2R_2$. In what follows we consider the dual transmitter ping-pong mode to illustrate the main form of governing equations.

In either case, when we form the complex product $s_1 s_2^*$ for interferometry we cancel the scattering phase terms (assuming they are the same; we shall see later how good this approximation is in practice) and keep the geometrical phase. In effect we now obtain a signal phase which depends only on the difference in range between the two positions, $\Delta R = R_1 - R_2$, as shown in equation (5.3):

$$\left. \begin{array}{l} s_1 = a_1 e^{-i\left(\frac{2\pi}{\lambda}2R_1 + \phi_{S1}\right)} \\ s_2 = a_2 e^{-i\left(\frac{2\pi}{\lambda}2R_2 + \phi_{S2}\right)} \end{array} \right\} \Rightarrow s_1 s_2^* = A e^{-i\left(\frac{4\pi}{\lambda}\Delta R\right)}$$
(5.3)

The key radar observable is now the interferometric phase ϕ , as defined in equation (5.4):

$$\phi = \arg\{s_1 s_2^*\} = -\frac{4\pi}{\lambda} \Delta R + 2\pi N \quad N = 0, \pm 1, \pm 2, \dots$$
 (5.4)

Note that this interferometric phase is 2π ambiguous; that is, for every shift of half a wavelength $\lambda/2$ in range difference ΔR the phase difference repeats itself. We therefore face problems if we wish to invert this equation and determine ΔR from a measurement of phase, as there are multiple solutions. This is called the 'phase unwrapping problem' (Bamler, 1998), and its solution requires the use of additional information, such as use of phase from nearby scattering points to generate spatial phase gradients and force continuity, or the use of multiple baselines (data from three or more spatial positions).

In what follows we consider three different types of radar interferometry used in remote sensing: across-track, differential, and along-track modes.



Fig. 5.1 Baseline geometry of radar interferometry

5.1.1 Across-track interferometry

The next key step is to recognize that from a measurement of phase and knowledge of the components of baseline B we can estimate the height of a scattering point h_0 (see Figure 5.1). This phase-to-height conversion was the original motivation for the development of radar interferometry (Graham, 1974). This conversion arises directly from the geometry of the triangle PS₁S₂ shown in Figure 5.1. From the cosine law applied to this triangle we can then derive the following relationship between the height at point P, h_0 , the system geometry, and the interferometric phase, as shown in equation (5.6):

$$\sin(\theta - \delta_b) = \frac{R_2^2 - R_1^2 - B^2}{2BR_1} \Rightarrow h_0 = H - R_1 \cos\theta$$
(5.6)

From this, knowing B (often to an accuracy of mm), the baseline angle δ_b , and the two ranges R_1 and R_2 , we can obtain an estimate of θ for the point P. This then enables calculation of height at the point P above some datum H, as shown in equation (5.6).

When direct measurements are made of range from two positions, this process is called radar stereo-grammetry. In this case, the accuracy of the final height estimate depends primarily on the accuracy of the range difference measurements. It is this range difference parameter that can be measured much more accurately using interferometry. This consequently leads to a much higher accuracy estimate of the angular position, and hence of the height of the scatterer. To see this, consider setting $R_2 = R_1 + \Delta R$ in equation (5.6), and assuming that $R_1 \gg B$ and $R_1 \gg \Delta R$, to obtain the following relationship between phase and height $z = h_0$.

$$B\sin(\theta - \delta_b) = \Delta R + \frac{\Delta R^2}{2R_1} - \frac{B^2}{2R_1} \approx \Delta R = \frac{\phi\lambda}{4\pi}$$

$$\Rightarrow \sin(\theta - \delta_b) = \frac{\lambda\phi}{4\pi B} \Rightarrow z = H - R\cos\theta$$
(5.7)

This gives the phase and position of a single scatterer. In the presence of a second scatterer we obtain a change of phase obtained from equation (5.7), as shown in equation (5.8):

$$\Delta \phi = \frac{4\pi}{\lambda} B \cos(\theta - \delta_b) \Delta \theta \\\Delta z = R \sin\theta \Delta \theta - \Delta R \cos\theta \}$$

$$\Rightarrow \Delta \phi = \frac{4\pi B \cos(\theta - \delta_b)}{\lambda R \tan\theta} \Delta R + \frac{4\pi B \cos(\theta - \delta_b)}{\lambda R \sin\theta} \Delta z \qquad (5.8)$$

$$= \frac{4\pi B_n}{\lambda R \tan\theta} \left(1 + \frac{\Delta z}{\Delta R} \frac{1}{\cos\theta} \right)$$

where B_n is the normal component of the baseline (see Figure 5.7). For scatterers lying in a plane, $\Delta z = 0$, and the phase gradient can then be related to the normal component of the baseline B_n as

$$\frac{\partial \phi}{\partial R} = \frac{4\pi B_n}{\lambda R \tan \theta}$$
(5.9)

This is called the flat-earth component of the phase, and can be used to either estimate the baseline from the phase gradient over flat terrain or, since generally



Fig. 5.2 Approximate ray-path geometry for phase estimation

a high-frequency component of the phase signal, it can be removed by the process of flat-earth removal (discussed in Section 5.1.1). When this is removed this leaves us with a simple linear relationship between phase and elevation Δz .

Another informative way to derive this relation—employing a local coordinate system around the point P, and which also leads to a linear relationship between phase and height of the form $\phi = \beta_z h_0$, where the scale factor β_z is called the interferometric wavenumber—is as follows. We first consider construction of a local coordinate system around the point P, as shown in Figure 5.2. The y axis represents the local surface tangent, and z the local vertical. Considering first the y axis, the extra contribution to phase for a point P, separated locally from the origin by distance y, is approximately ysin θ , as shown in Figure 5.2.

Using this we can then write the phase for signals scattered from point P at ends 1 and 2 of a baseline, as shown in equation (5.10a):

$$\left.\begin{array}{l} s_1 = a_1 e^{-i\frac{4\pi}{\lambda}(R+y\sin\theta_1)} \\ s_2 = a_2 e^{-i\frac{4\pi}{\lambda}(R+y\sin\theta_2)} \end{array}\right\} \Rightarrow s_1 s_2^* = A e^{i\frac{4\pi}{\lambda}(\sin\theta_2 - \sin\theta_1)} \tag{5.10a}$$

$$s_1 s_2^* = A e^{i\frac{4\pi}{\lambda}y(\sin\theta_2 - \sin\theta_1)} = A e^{i\frac{4\pi}{\lambda}2y\cos(\frac{\theta_1 + \theta_2}{2})\sin(\frac{\theta_2 - \theta_1}{2})} \approx A e^{i\frac{4\pi\Delta\theta}{\lambda}y\cos(\theta)}$$
(5.10b)

$$s_1 s_2^* = A e^{-i\frac{4\pi}{\lambda} z(\cos\theta_2 - \cos\theta_1)} = A e^{i\frac{4\pi}{\lambda} 2z\sin(\frac{\theta_1 + \theta_2}{2})\sin(\frac{\theta_2 - \theta_1}{2})} \approx A e^{i\frac{4\pi\Delta\theta}{\lambda} z\sin(\theta)}$$
(5.10c)

In case the angular difference to the point $\Delta \theta = (\theta_2 - \theta_1)$ is small (that is, B \ll R), we can further simplify this expression, as shown in equation (5.10b). Similarly, for points P shifted in the z direction (the right-hand side of Figure 5.2) we have an extra phase term, $-z \cos \theta$, which leads to a corresponding interferometric phase, as shown in equation (5.10c). Combining these two, we obtain the following expression for the total interferometric phase as a function of y and z around the point P.

$$\phi_{if}(y,z) = \frac{4\pi\,\Delta\theta}{\lambda}\,(y\cos\theta + z\sin\theta) \tag{5.11}$$

Currently we see that the phase varies not only with height of the point P (the coordinate z) but also for shifts in surface position (y). This somewhat disturbs our desire to have a clean phase-to-height conversion. We can remove this problem and obtain our desired result by shifting the frequency of the signal collected at position 2 *before* forming the interferogram, using a process called range spectral filtering, as follows.

5.1.1.1 Range spectral filtering

One way to remove the surface (y coordinate) phase dependence in equation (5.11) is to generate an interferogram, for each frequency f in the spectrum of signal 1, using a *shifted* frequency component of signal 2, as shown in equation (5.12) (Gatelli, 1994):

$$\phi = \arg(s_1(f)s_2^*(f + \Delta f)) \tag{5.12}$$

In order to completely cancel the y-dependence of phase in equation (5.11) for arbitrary position P, we must employ a shift Δf based on geometry $\Delta \theta$, as shown in equation (5.13). This, then, yields a scale factor β_z equal to the desired vertical sensitivity of the interferometer: namely, $\frac{\partial \phi}{\partial z}$.

$$if \Delta f = \frac{\Delta \theta}{\tan \theta} f$$

$$\Rightarrow \Delta \beta = \beta \frac{2\Delta \theta}{\tan \theta} = \frac{4\pi \Delta \theta}{\lambda \tan \theta}$$

$$\Rightarrow \phi = \phi_{if} + \Delta \beta (z \cos \theta - y \sin \theta)$$

$$\Rightarrow \phi = \frac{4\pi \Delta \theta}{\lambda} (y \cos \theta + z \sin \theta) + \frac{4\pi \Delta \theta}{\lambda \tan \theta} (z \cos \theta - y \sin \theta)$$

$$\Rightarrow \phi = \frac{4\pi \Delta \theta}{\lambda \sin \theta} z = \frac{\partial \phi}{\partial z} z = \beta_z z$$

$$\Rightarrow \beta_z = \frac{4\pi \Delta \theta}{\lambda \sin \theta}$$
(5.13)



Fig. 5.3 Wavenumber space geometrical interpretation of radar interferometry

This transformation can be given a simple geometrical interpretation if we employ a two-dimensional wavenumber representation of interferometry as shown in Figure 5.3. Here we show the complex signal from the first sensor position as a vector in the β_y, β_z plane with magnitude equal to $2\pi/\lambda$ and angle θ to the point P. The second position (before frequency shift) then yields a vector with the same length but rotated about the origin by $\Delta\theta$, as shown. If we consider the y components only, then in order to give s_2 the same projection on the y axis as s_1 requires a radial shift in wavenumber by $\Delta\beta$ (which is negative if $\Delta\theta$ is positive) before forming the interferogram.

Furthermore, this idea can be generalized to the bistatic surface scattering case, using a concept known as the memory line, first described in Le (1998). This can be motivated by considering the geometrical representation of a general bistatic scattering geometry in wavenumber space, as shown in Figure 5.4. Here we see that a bistatic angle ψ leads to a shift of the point representation in wavespace in both angle and radial coordinates. To see this, just consider the general



Fig. 5.4 Geometry of generalised bistatic radar in real and wavenumber domains

phase term as shown in Figure 5.5. Now consider forming an interferogram between signals for two different bistatic geometries with corresponding wavespace coordinates r_1,ϕ_1 and r_2,ϕ_2 . In order that the phase difference between signals has no y-dependence and depends only on height, then equal projections onto the y axis are required, which in turn requires that equation (5.14) holds:

$$|\beta|\cos\frac{\psi}{2}\cos\left(\theta + \frac{\psi}{2}\right) = C \tag{5.14}$$

Consequently, we can now change any of the three parameters, θ , ψ and β , but as long as the product shown in equation (5.14) remains constant then the interferometric phase will be a function only of height. Equation (5.14) is the equation of a line—the so-called memory line of the interferometer—and all systems lying along this line satisfy our desired phase dependency. We see that for the special case of backscatter ($\psi = 0$) this constraint reduces to $\beta \cos \theta$ being constant, which is the origin of the frequency shift introduced in equation (5.13).

Returning now to the case of backscatter, we can obtain a useful approximation for $\Delta\theta$ as found in equation (5.13) in the special case $R \gg B$ (as occurs for spaceborne radar for example). In Figure 5.6 we show how the angular baseline width $\Delta\theta$ can be approximated from the geometry as $\Delta\theta \approx B_{\perp}/R$. Note that this expression does not include the absolute baseline B, but only its component perpendicular to the range line. This leads us to consider the various spatial components of the baseline, as follows.

5.1.1.2 Baseline components

It is often convenient to perform a geometric decomposition of the baseline vector as shown in Figure 5.7, which includes various important components. The horizontal (B_H) and vertical (B_V) can be simply defined in terms of the baseline orientation δ_b , as shown, while the parallel and perpendicular components require knowledge of both the baseline angle and the mean angle of incidence. The main conclusion from Figures 5.6 and 5.7 is that we can often approximate the vertical sensitivity of an interferometer in terms of the baseline-to-wavelength ratio, as shown in equation (5.15):

$$\beta_z = \frac{4\pi \,\Delta\theta}{\lambda \sin \theta} \approx \frac{4\pi B_\perp}{\lambda R_0 \sin \theta} = \frac{4\pi \cos(\theta - \delta_b)}{R_0 \sin \theta} \frac{B}{\lambda}$$
(5.15)

We can therefore adjust the sensitivity of the interferometer by adjusting the B/λ ratio. This provides us with the ability to 'tune' the performance of the





Fig. 5.5 Bistatic triangle construction for radar interferometry



Fig. 5.6 Baseline geometry triangle

sensor to different applications—a feature not available in radar polarimetry, and one which we shall see is very important in the development of polarimetric interferometry.

An important measure of the sensitivity of the interferometer is the height corresponding to a phase shift of 180° —the π -height. This can be derived from equation (5.15), as shown in equation (5.16):

$$h_{\pi} = \frac{\pi}{\beta_z} = \frac{\lambda \sin \theta}{4\Delta \theta} \approx \frac{\lambda}{4} \frac{R_0 \sin \theta}{B_{\perp}}$$
(5.16)

Note that since $R_0 \gg B$, this height is much greater than a wavelength. We shall see later how in differential interferometry we can obtain much higher sensitivities when the π -height itself reduces to $\lambda/4$. Before leaving this topic we highlight three important issues.

1. The frequency shift used in equation (5.12) is actually a function of the *local* angle of incidence at the point P. Hence, in the presence of topographic variations it is affected by range slopes. For example, if the terrain is sloped towards the radar then the effective angle of incidence required to determine the frequency shift is reduced. In the extreme case, when the surface slope equals the radar angle of incidence we ultimately determine effective normal incidence onto the surface and lose sensitivity completely. In this case we obtain so-called 'blind' angles for the interferometer. In general, therefore, in the presence of range slope η we must modify the expression for spectral shift, as shown in equation (5.17):

$$\Delta f = -\frac{cB_{\perp}}{R_0\lambda\tan\left(\theta - \eta\right)} \tag{5.17}$$

As an example we consider evaluation of the required spectral shift for typical values of a space radar (ERS-1) operating at 23° of incidence at C band ($\lambda = 5.66$ cm) and transmitting a signal with bandwidth 15.5 MHz with a slant range of 840 km. Figure 5.8 shows the required frequency shift as a function of surface slope. We note the following important features:

- (a) As slope increases from zero, the required frequency shift eventually exceeds the bandwidth of the pulse $(\Delta f / W = 1)$. This is the start of a band of 'blind angles', and the width of this band is defined from this point until the slope increases to the extent that the frequency shift returns to the pulse bandwidth. This high angle range, however, is plagued by layover distortion in radar imaging, whereby points at the top of a slope are at closer range to the radar than those at the foot of the slope.
- (b) In the negative range slope direction (away from the radar) the required shift slowly decreases to zero for grazing incidence on the surface. Thereafter the slope disappears into shadow, and there is no longer a measurable radar backscatter return.

Correct application of the frequency shift concept therefore requires detailed knowledge of the terrain slope. This can be obtained, for example, if a reference digital elevation model (DEM) is available, from which the range slopes can be calculated by spatial differentiation. In this way an adaptive variation of spectral shift can be performed. In the absence of such topographic information,



Fig. 5.8 Example of frequency shift versus surface slope

Fig. 5.9 Schematic representation of range spectral shift

however, a compromize shift corresponding to the mean slope can be used, recognising that there will be some errors in the presence of severe topography.

5.1.1.3 Critical baseline

We see from equation (5.15) that we can always increase the sensitivity of the interferometer by increasing the baseline. There is, however, a limit to this process, as eventually the frequency shift required will exceed the available bandwidth of the radar signal. If this bandwidth is *W* Hz, then by equating the shift to *W* we can obtain an expression for the critical baseline, as follows:

$$W = -\frac{cB_{\perp,crit}}{R_0\lambda\tan\left(\theta - \eta\right)} \Rightarrow B_{\perp,crit} = \left|\frac{\lambda WR_0\tan\left(\theta - \eta\right)}{c}\right|$$
(5.18)

from which we see that it is the product of bandwidth times wavelength λW that is important in setting the maximum sensitivity of the interferometer. Thus for a given geometry, a system with 10 MHz of bandwidth at 600 MHz (0.5-m wavelength) has the same critical baseline as a 100-MHz bandwidth system at 6 GHz.

Note also that this frequency shift leads to a reduction in useful bandwidth and hence to a reduction in range resolution of the radar system. This arises because with a shift there is a reduced common overlap between the spectra at ends 1 and 2 of the baseline, as shown schematically in Figure 5.9. The slant range resolution of an interferometer Δr , normally equal to c/2W for a single radar sensor, then has the modified form shown in equation (5.19):

$$\Delta r = \frac{c}{2(W - \Delta f)} \tag{5.19}$$

5.1.1.4 Flat earth removal

Returning to the general expression for interferometric phase (equation (5.8)), one parameter of importance is the fringe frequency with slant range f_s , defined again in equation (5.20):

$$\begin{split} \phi &= 2\beta \Delta R \approx 2\beta B_{||} = 2\beta B \sin \theta \\ z &= h - R \cos \theta \end{split} \Longrightarrow \begin{array}{l} \Delta \phi &= 2\beta B \cos \theta \Delta \theta \\ \Delta z &= R_0 \sin \theta \Delta \theta - \Delta R \cos \theta \end{aligned}$$
$$\to \Delta \phi &= \frac{4\pi B_\perp}{\lambda R_0 \tan \theta} \Delta R + \frac{4\pi B_\perp}{\lambda R_0 \sin \theta} \Delta z \qquad (5.20)$$
$$\to f_s &= \frac{1}{2\pi} \frac{\partial \phi}{\partial R} \bigg|_{\Delta z = 0} = \frac{2B_\perp}{\lambda R_0 \tan \theta} \end{split}$$

This is interpreted physically as the rate of change of phase across a surface for which $\Delta z = 0$ —a 'flat earth'. For example, for a space radar operating at a 23 degree angle of incidence at C-band (5.66=cm wavelength) at a slant range of 840 km, and with a perpendicular baseline of 200 m, this leads to a phase rate of 1 cycle/50 m (which is only around five slant range cells for a 15-MHz bandwidth typical of space radars). This high-frequency phase term is therefore often removed at the start of interferometric analyses, so generating a constant phase for flat terrain, which then helps emphasize any topographic variations that may be present. It also facilitates coherence estimation, as we show in the next section. It can always be returned to the phase signal at the end of processing.

Flat-earth processing involves multiplication of the interferogram by a conjugate phase signal to cancel the flat-earth components, as shown in equation (5.21). Here *r* is the slant range coordinate of the radar.

$$s_1 s_2^* e^{-i2\pi f_s r}$$
 (5.21)

This is called 'flat earth' removal for interferometric processing. Note that it is not a replacement for spectral shift processing, and its sole purpose is to remove the high-frequency modulation of the underlying phase to facilitate further processing and interpretation. Note that for airborne radar or other large swath geometries, when the angle of incidence varies across the range swath, the spatial frequency of the flat-earth signal is no longer constant but increases with decreasing angle of incidence, leading to a 'chirp' or variable frequency reference. In this case the same procedure can be used for flat-earth removal, but f_s must be evaluated more carefully by accurate evaluation of the $\Delta\theta$ term.

We now briefly consider two important variations on radar interferometer design. The first—called 'differential interferometry'—is designed to maximize sensitivity to changes in repeat-pass sensors; while the second—along-track interferometry (ATI)—is designed to sense the velocity rather than position vector of scatterers.



Fig. 5.10 Geometry of differential interferometry

5.1.2 Introduction to differential interferometry

An important special case of radar interferometry arises for zero spatial baselines, B = 0, but non-zero temporal baseline T_B . In this case, from equation (5.3) we expect the interferometric phase to be zero for all surfaces. However, this ignores any motion of the surface that may have occurred between passes of the sensor. Figure 5.10 shows a schematic representation of this situation. In the solid line is shown a surface measured from position Q at time t_1 to obtain s_1 . The dashed line is the new position of the surface at time $t_2 = t_1 + T_B$ when measured again by the sensor at Q to obtain s_2 . The interferogram formed as $s_1 s_2^*$ will have non-zero phase because of the radial component of shift of the surface. If different parts of the surface move by different amounts then we obtain a phase map of these displacements in the interferogram. Equation (5.22) shows how such an interferometer measures the projection of the surface displacement vector <u>d</u> onto the line of sight.

$$\phi = \frac{4\pi}{\lambda} \left(R_2 - R_1 \right) = \frac{4\pi}{\lambda} \left| \underline{d} \right| \sin \left(\theta - \delta_d \right)$$
(5.22)

Hence we see that there is no sensitivity to movement perpendicular to the line of sight (when $\theta = \delta_d$). Note that in this case there is no baseline scaling and the sensitivity is maximum, given by the full wavenumber β . If we define the quantity $d = |\underline{d}| \sin (\theta - \delta_d)$ as an equivalent height to the offset baseline case, then the π -displacement is given as

$$d_{\pi} = \frac{\lambda}{4} \tag{5.23}$$

If the wavelength is a few centimetres and the phase accuracy is around 5° , then this can lead to sensitivities of the order of mm. There are, however, two major problems with differential interferometry:

The first is difficulty in obtaining a true zero spatial baseline (exact repeatpass). Hence in practice there is always some spatial baseline B to consider, which inevitably contributes a phase component sensitive to topography, as considered earlier. Following flat-earth removal, the total interferometric phase can therefore be more accurately written, as shown in equation (5.24):

$$\phi = \phi_{topo} + \phi_{displacement} = -\frac{4\pi B_{\perp}z}{\lambda R_0 \sin \theta} + \frac{4\pi}{\lambda}d$$

$$\Rightarrow \phi_{displacement} = \phi - \phi_{topo}$$
(5.24)

There are two ways to combat this topographic dependency. The first is to use a reference DEM to estimate the topographic phase component, and to then remove its contribution from the phase signal, as shown in the lower portion of equation (5.24). One potential problem with this approach is that the reference DEM employed will have some errors, and these can propagate into the displacement estimate. For example, if the DEM has an elevation error Δz then this translates into an equivalent phase error of $\phi_{\Delta z}$ which in turn is interpreted as a ground motion $d_{\Delta z}$, as shown in equation (5.25):

$$\phi_{\Delta z} = \frac{4\pi B_{\perp}}{\lambda R_0 \sin \theta} \Delta z \Rightarrow d_{\Delta z} = \frac{\lambda}{4\pi} \phi_{\Delta z} = \frac{B_{\perp}}{R_0 \sin \theta} \Delta z \qquad (5.25)$$

This can be expressed more conveniently in terms of the critical baseline for a sensor with bandwidth W operating at centre frequency f_0 , as shown in equation (5.26):

$$d_{\Delta z} = \frac{B_{\perp}}{R\sin\theta} \Delta z = \frac{B_{\perp}}{B_{crit}} \frac{W}{f_0} \frac{\Delta z}{\cos\theta}$$
(5.26)

Clearly the error can be minimized by reducing the baseline, and in the limit of zero spatial baseline the topographic error is removed. The second approach (when a reference DEM is not available) is to employ three (or more) passes to eliminate the topographic phase dependence. If $n \ge 3$ passes are available then we can define a set of interferometric parameters, as shown in equation (5.27):

- time baselines $\Delta t_{n-m} = t_n - t_m$ - spatial baselines $B_{\perp,n-m}$ (5.27) - interferometric phases ϕ_{n-m}

We can then decompose the set of phases ϕ_{n-m} into a part proportional to the spatial baselines (topography) and a part proportional to the temporal baselines (displacement at constant velocity). To illustrate this we consider the simplest case of three passes, yielding complex signals s_1 , s_2 and s_3 . The first signal, s_1 , is then used as a 'master' to generate interferograms with both s_2 and s_3 . The first baseline B_{12} we wish to be dominated by topographic effects, and so should be as large as possible, with a small temporal baseline so as to minimize displacement effects. The second baseline, B_{13} , however, should be dominated by temporal effects, and so should combine a small spatial with long temporal baselines. With this combination the displacement phase can be directly estimated as shown in equation (5.28). (Note that if the baseline ratio is too large (>4) then phase unwrapping may be required of the ϕ_{12} interferometric phase before scaling.)

$$\phi_{displacement} = \phi_{13} - \frac{B_{13}}{B_{12}}\phi_{12} \tag{5.28}$$

The second problem faced in differential interferometry is the effect of wave propagation between the sensor and the surface. The propagation of microwaves through the atmosphere causes a phase shift due to variations in refractive index. For repeat-pass sensors the changing atmosphere causes a change in this phase and hence a phase error in the interferogram. Such propagation effects can be large, for example, for a spaceborne radar at C-band, as atmospheric delays can cause an error of half a fringe (π radians), or in extreme cases up to three

fringes. For low-frequency radars, phase shifts due to propagation through the ionosphere can cause similar problems (Freeman, 2004). In general, therefore, the phase of an interferogram can be written in component form, as shown in equation (5.29):

$$\phi = \phi_{flat} + \phi_{topo} + \phi_{displacement} + \phi_{propagation}$$
$$= -\frac{4\pi B_{\perp} r}{\lambda R_0 \tan \theta} - \frac{4\pi B_{\perp} z}{\lambda R_0 \sin \theta} + \frac{4\pi}{\lambda} d + \phi_{propagation}$$
(5.29)

Note that the propagation phase does not depend on baseline and hence cannot be removed by baseline diversity. It is embedded as an error source in the displacement phase. Recently there have been several techniques proposed for separating the propagation from displacement phase by employing the former's distinct lack of temporal correlation combined with high spatial correlation arising from the fractal nature of the underlying atmospheric phase screen (Kampes, 2006). However, this method requires the acquisition of a large number of passes, and hence takes us beyond the bounds of a basic introduction. Instead we turn to consider the third important type of interferometer: ATI.

5.1.3 Along track interferometry (ATI)

The third important interferometric configuration to be considered is along track interferometry, or ATI. This is a single-pass configuration with two radar systems displaced with a spatial baseline *parallel* to the direction of motion of the platform, as shown schematically in Figure 5.11. The key idea is that in this configuration the spatial baseline and platform velocity combine to obtain a short temporal baseline Δt (typically of the order of 10–100 msecs), during which the scatterer (moving itself with velocity *v*) will move, and thus cause a change in range, which leads to a phase shift.

We can then quantify the relationship between interferometric phase and velocity as shown in equation (5.30):

$$\phi = \frac{4\pi}{\lambda} \Delta R = \frac{4\pi}{\lambda} \frac{\partial R}{\partial t} \Delta t = \frac{4\pi}{\lambda} v_{LOS} \frac{B}{v_r}$$
(5.30)

where v_{LOS} is the line-of-sight component of the velocity vector \underline{v} of the point P. Hence ATI remains blind to velocities parallel to the platform motion. Such a technique can be used to measure ocean currents and glacier motion, as well as the speed of point scatterers such as ships and land vehicles. One key limitation of this idea is the maximum temporal baseline that can be used. Decorrelation effects in the scatterer eventually lead to a loss of coherence (discussed in the next section). For this reason the temporal baseline needs to be designed with a measure of the typical scatterer decorrelation time in mind. This brings us to consider decorrelation and its relation to an important new observable, interferometric coherence.

5.2 Sources of interferometric decorrelation

In the previous section we showed how interferometric phase can be related to several important surface parameters (height, velocity, displacement, and so



Fig. 5.11 Geometry of along-track interferometry

on), depending on the configuration used. However, so far we have ignored the influence of noise and its impact on phase estimation. In polarimetry we saw that noise arises from depolarisation and is manifest as an increase in scattering entropy. In this section we first consider a formalism to include noise effects in radar interferometry, and then consider a set of various potential sources of noise (Zebker, 1992). Some of these are system related (signal-to-noise ratio, for example), but others are related to wave scattering effects (volume and baseline decorrelation in particular) and hence can be considered analogous to wave depolarisation effects in polarimetry. By considering such coherent scattering in detail, we will see how we can then turn the noise problem around and use the interferometric coherence as a new radar observable to help estimate surface and volume scattering parameters. This will then lead us to consider, in the next chapter, combinations of polarimetry with interferometry.

We start with a general expression for interferometric phase ϕ as the sum of 'signal terms' ϕ_{if} and a noise term ϕ_n , (equation (5.31)) characterized by its statistical moments. Generally the noise term will have zero mean, but is characterized by non-zero standard deviation σ_{ϕ} .

$$\phi = \phi_{if} + \phi_n \tag{5.31}$$

To see the impact of such stochastic fluctuations in the phase on surface parameters, consider the important special case of surface height estimation using across-track radar interferometry. As shown in equation (5.32), the phase variance σ_{ϕ} leads to a scaled height variance, derived from the relation of phase to changes in slant range $\sigma_{\Delta R}$ and then using equation (5.15) to relate range to height via the normal baseline.

$$R = \frac{\lambda}{4\pi}\phi \Rightarrow \sigma_{\Delta R} = \frac{\lambda}{4\pi}\sigma_{\phi} \Rightarrow \sigma_{h} \approx \frac{R_{0}\sin\theta}{B_{\perp}}\sigma_{\Delta R} \approx \frac{R_{0}\sin\theta}{B_{\perp}}\frac{\lambda}{4\pi}\sigma_{\phi} \quad (5.32)$$

This relation can be used to estimate errors in surface height based on system parameters (baseline geometry and angle of incidence) and the phase variance. To proceed, we need to further investigate the different ways in which phase noise can be generated in radar interferometry. To do this we first relate noise variance to an underlying coherence.

We start by employing a coherency matrix formulation of radar interferometry, as shown in equation (5.33). Here again we can define a useful secondary parameter: the interferometric coherence, with a magnitude between 0 (pure noise) and 1 (pure signal).

$$[T_2] = \begin{bmatrix} \langle |s_1|^2 \rangle & \langle s_1 s_2^* \rangle \\ \langle s_2 s_1^* \rangle & \langle |s_2|^2 \rangle \end{bmatrix} \Rightarrow \tilde{\gamma} = \frac{\langle s_1 s_2^* \rangle}{\sqrt{\langle |s_1|^2 \rangle \langle |s_2|^2 \rangle}}$$
(5.33)

From Appendix 3 it then follows that the phase variance can be related to the coherence by the following Cramer–Rao bounds (Seymour, 1994):

$$\sigma_{\phi} \le \sqrt{\frac{1 - |\gamma|^2}{2L |\gamma|^2}} \quad \sigma_{|\gamma|} \le \frac{1 - |\gamma|^2}{\sqrt{2L}} \tag{5.34}$$

where *L* is the number of independent samples used in forming the average <..>. Note that often the coherence itself is estimated from the data (Touzi, 1999), in which case it has an estimation variance defined using the Cramer–Rao bound also shown in equation (5.34). We now take a closer look at the origin of these fluctuations. We look at four factors: signal-to-noise ratio, temporal decorrelation, baseline, and volume decorrelation.

5.2.1 Signal-to-noise decorrelation

The two complex signals s_1 and s_2 can first be decomposed into signal (*a*) and noise (*n*) terms, as shown in equation (5.35):

$$s_1 = a + n_1$$

 $s_2 = a + n_2$
(5.35)

Now we must invoke some assumptions about the statistical distribution of the noise terms. The most common assumption, based on the central limit theorem, is that the noise terms are complex Gaussian random variables and hence have uniform phase distributions and are uncorrelated both with the signal (a) and with each other. Under this assumption the coherence can be evaluated as shown in equation (5.36):

$$\gamma_{snr} = \frac{|a|^2}{|a|^2 + |n|^2} = \frac{SNR}{1 + SNR}$$
(5.36)

Here we see a simple relation between the signal-to-noise ratio (SNR) and coherence. As the SNR tends to infinity (zero noise) then the coherence tends to unity, while if the SNR tends to zero then the coherence also tends to zero. Figure 5.12 shows how the coherence is related to SNR (expressed in dB).



Fig. 5.12 Relationship between signal-tonoise ratio and coherence

5.2.2 Temporal decorrelation

A second key model employed for a noise source is to assume that the noise has constant amplitude but a Gaussian distribution of phase. This model is more appropriate when we have a signal in the presence of an unwanted 'clutter' background. In this case we can use the following identity for the average phase difference between stochastic signals with Gaussian phase statistics:

$$\begin{cases} s_1 = a + me^{i\phi_1} \\ s_2 = a + me^{i\phi_2} \end{cases} \Rightarrow \left\langle e^{-i(\phi_1 - \phi_2)} \right\rangle \approx e^{-\frac{\sigma_\phi^2}{2}}$$
(5.37)

This then enables us to calculate an expression for the coherence, as shown in equation (5.38):

$$\gamma = \frac{|a|^2 + |m|^2 e^{-\sigma_{\phi}^2}}{|a|^2 + |m|^2} = \frac{\text{SCR} + e^{-\sigma_{\phi}^2}}{\text{SCR} + 1}$$
(5.38)

where SCR is now the signal-to-clutter ratio. The most important application of this model is to temporal decorrelation in repeat-pass interferometry. In this case the 'clutter' noise is caused by motion of scatterers (such as wind-driven vegetation) between passes. If the rms motion along the line of sight is $\delta_{\rm rms}$, then the phase variance and hence coherence in equation (5.38) can be simply related to this shift, as shown in equation (5.39):

$$\sigma_{\phi} = \frac{4\pi}{\lambda} \delta_{rms} \Rightarrow \gamma = \frac{SCR + e^{-\frac{16\pi^2}{\lambda^2} \delta_{rms}}}{SCR + 1}$$
(5.39)

Importantly, we see that this coherence depends on the ratio of rms motion to wavelength, and hence for a given shift the effect on coherence is worse for higher frequencies. This drives us to consider lower frequencies to minimize the effects of temporal decorrelation in repeat-pass interferometry (Hagberg, 1995; Askne, 1997, 2003, 2007).

Note that in the general case we can have a combination of these statistical effects, such as temporal decorrelation γ_t in combination with noise decorrelation γ_{snr} . In this case the coherence is formed from products of triple sums, as shown in equation (5.40). The most important consequence of this is that coherence always decomposes in a multiplicative series of component terms (to be contrasted with polarimetric decomposition which led to expansion as a sum of component terms).

$$s_{1} = a + me^{i\phi_{1}} + n_{1} \\ s_{2} = a + me^{i\phi_{2}} + n_{2} \end{cases} \Rightarrow \gamma = \frac{|a|^{2} + |m|^{2} e^{-\sigma_{\phi}^{2}}}{|a|^{2} + |m|^{2} + |n|^{2}} \\ = \frac{|a|^{2} + |m|^{2} e^{-\sigma_{\phi}^{2}}}{|a|^{2} + |m|^{2}} \cdot \frac{|a|^{2} + |m|^{2}}{|a|^{2} + |m|^{2} + |n|^{2}} = \gamma_{l}\gamma_{snr}$$
(5.40)

Although only shown for a combination of two components in equation (5.40), the same argument can be used for coherence of an arbitrary mixture of independent terms. For example, in addition to SNR and temporal effects there are always some additional sources of coherence loss due to processing errors.

Typically in radar applications the two signals s_1 and s_2 are formed from coregistered synthetic aperture radar (SAR) images (often collected at different times of a repeat orbit), and in practice it is impossible to exactly match the two radar signals (see Chapter 9). There will always be some small residual fractional offset in range and azimuth pixel size δ_{rg} and δ_{az} between the two images (Krieger, 2005). This causes a coherence loss component given by equation (5.41):

$$\gamma_{proc} = \frac{\sin \pi \,\delta_{rg}}{\pi \,\delta_{rg}} \frac{\sin \pi \,\delta_{az}}{\pi \,\delta_{az}} \tag{5.41}$$

Current processing accuracies are limited to about 1/10 of a pixel in both range and azimuth, and so we see that this error is independent of baseline and has a value around 0.97. This error becomes particularly significant in the case of small spatial baselines, where it can become the dominant source of decorrelation. To combine this error source with the other two we simply extend the decomposition of equation (5.40), as shown in equation (5.42):

$$\gamma = \gamma_{SNR} \gamma_t \gamma_{proc} \tag{5.42}$$

This approach gives us the ability to consider different independent decorrelation sources and include them in the final expression for coherence in a straightforward (multiplicative) way. In particular, there are two more important scattering-based decorrelation sources to be considered: baseline and volume decorrelation, which we now consider in turn.

5.2.3 Baseline decorrelation

Surface scattering can give rise to an important source of coherence loss termed baseline decorrelation (Zebker, 1992; Gatelli, 1994). The origin of this process can be found in the ideas of frequency shift, as discussed in equation (5.13). There we showed that the expression for interferometric phase (before spectral shift filtering) contains a dependence on the y or surface coordinate of the scattering point. Therefore, if we have a distribution of scattering points within a range cell they will add coherently to yield some resultant complex return. However, when we shift position to the other end of the baseline we obtain a slightly different coherent sum from the same set of points (simply because the surface component of the wave vector has changed). This fluctuation in the complex sum for surface scatterers leads to a loss of coherence, as we now show.

We can immediately see one important additional benefit of performing the spectral shift *before* interferogram formation. If the spectral shift is applied, then by definition the contributions from both ends of the baseline have the same surface component of wavenumber and hence the same coherent phase addition for surface scatterers. Following spectral filtering the coherence equals 1, and baseline decorrelation is removed. From our discussion around equation (5.18) we see that this will be possible up to a maximum baseline, called the critical baseline, after which the spectral overlap will be zero and we obtain zero coherence. Thus the baseline decorrelation is given by the ratio of shifted spectral overlap to total bandwidth W. This results in the expression for baseline



$$\gamma_B = \frac{B_{crit} - B_{\perp}}{B_{crit}} = 1 - \frac{B_{\perp}}{B_{crit}} = 1 - \frac{cB_{\perp}}{W\lambda R_0 \tan(\theta - \eta)}$$
(5.43)

We reiterate that this decorrelation occurs only if spectral filtering is not applied. By employing a spectral shift we can always ensure that $\gamma_B = 1$ (up to a maximum separation of the critical baseline, although note from equation (5.19) that the price to pay for this shift is that the range resolution reduces).

Before leaving this topic, we note one important scenario that always generates unit baseline coherence, independent of spatial baseline and spectral shift. This is when the resolution cell contains only a single point scatterer. To see this, consider an alternative interpretation of critical baseline in terms of an effective scattering diagram, as shown schematically in Figure 5.13. Here we show a surface resolution element, which for a distributed surface scatterer (shown in grey) has a spatial extent bounded by the bandwidth of the radar pulse W. This spatial segment has an apparent projected size Δ_{\perp} perpendicular to the line of sight, as shown in Figure 5.13. This projected surface element radiates back to the radar (the process of scattering), and has an effective beam width given by $\Delta \theta$, as shown. The critical baseline then occurs when this beamwidth fails to enclose *both* points 1 and 2 of the baseline. However, for a point scatterer (shown as the black disc), the spatial extent is not governed by the bandwidth but by the spatial size of the scatterer. In the limit of a point target Δ_{\perp} is a delta function, and hence $\Delta \theta$ becomes very large. The wide beamwidth therefore encloses all pairings of baseline end points 1 and 2, the critical baseline tends to infinity, and there is zero baseline decorrelation. This observation leads to the permanent scatterer (PS) technique in radar interferometry (Kampes, 2006), where high-accuracy positional information can be obtained from radar interferometry by restricting attention to point targets only-such as occur in urban areas, where there are many point-like man-made structures (see Chapter 9).

5.2.4 Volume decorrelation: the Fourier–Legendre series

In the previous section we saw how a random distribution of scatterers in a surface plane can cause decorrelation and loss of interferometric coherence through baseline (also called geometric) decorrelation. However, by employing range spectral filtering over terrain with known surface slope, we are always able to remove this decorrelation source (up to a limit given by the critical baseline). In a similar manner we note that a *vertical* distribution of scatterers will also cause a loss of coherence. This is termed volume decorrelation, as it often originates from volume scattering by layers of vegetation or snow/ice above the surface (Hagberg, 1995; Treuhaft, 1996).

However, one key distinguishing feature of volume decorrelation is that it is not possible to remove its effect by range spectral filtering. We saw from equation (5.13) that we can always choose Δk to remove the y but not the z dependence of interferometric phase. Therefore, two scatterers separated by a distance z will always have a phase difference given by the vertical wavenumber β_z , as shown in equation (5.44):

$$\phi = \beta_z z = \frac{4\pi \,\Delta\theta}{\lambda \sin \theta} z \approx \frac{4\pi B_\perp}{\lambda R_0 \sin \theta} z \tag{5.44}$$



Fig. 5.13 Geometric interpretation of baseline decorrelation

If we have a general variation of scattered power with z given by a vertical structure function f(z), the lower bound of which is at $z = z_o$, and the upper bound of which is at $z = z_0 + h_v$, where h_v is the height of the layer, then the interferometer will see a complex signal given by the weighted sum of contributions, as shown in equation (5.45), from which we can obtain an expression for the interferometric coherence, as shown in equation (5.46):

$$s_{1}s_{2}^{*} = \int_{z_{o}}^{z_{o}+h_{v}} f(z)e^{i\beta_{z}z}dz \xrightarrow{z'=z-z_{0}} e^{i\beta_{z}z_{o}} \int_{0}^{h_{v}} f(z')e^{i\beta_{z}z'}dz' \quad (5.45)$$
$$\tilde{\gamma} = e^{i\beta_{z}z_{o}} \frac{\int_{0}^{h_{v}} f(z')e^{i\beta_{z}z'}dz'}{\int_{0}^{h_{v}} f(z')dz'} = e^{i\beta_{z}z_{o}} |\tilde{\gamma}| e^{i\arg(\tilde{\gamma})} \quad (5.46)$$

Note that this is a complex coherence; that is, it has phase as well as magnitude, and part of the phase arises from the integral of the structure function f(z) shown in the numerator. This real non-negative function allows for arbitrary profile of scattering between the bottom and top of the layer (Cloude, 2006b). This relation shows that there is a direct relationship between the observed coherence and vertical structure properties of the scattering layer. For example, the height of the layer is found in the limits of the integral, the phase of the surface, while not equal to the phase of the coherence, is contained therein, and finally the structure function f(z) influences the coherence in both amplitude and phase.

Special cases of the structure function are often used in practice: for example, constant scattering amplitude or an exponential to more accurately model wave extinction effects in the layer (Treuhaft, 1996). Here we first develop a general theory of volume decorrelation based on arbitrary structure functions, and then specialize our discussion to these important special cases. The approach we use is to expand the bounded function f(z) in a Fourier–Legendre series, as follows. We first normalize the range of the integral in the numerator by a further change of variable, as shown in equation (5.47):

$$\int_{0}^{h_{v}} f(z') e^{i\beta_{z}z'} dz' \xrightarrow{z_{L} = \frac{2z'}{h_{v}} - 1} \int_{-1}^{1} f(z_{L}) e^{i\beta_{z}z_{L}} dz_{L}$$
(5.47)

We then rescale variation of the real non-negative function f(z) so that if $0 \le f(z) \le \infty$ then $f(z_L) = f(z) - 1$ and $-1 \le f(z_L) \le \infty$. Critically, we can now develop $f(z_L)$ in a Fourier–Legendre series on [-1,1], as shown in equation (5.48):

$$f(z_L) = \sum_n a_n P_n(z_L)$$

$$a_n = \frac{2n+1}{2} \int_{-1}^1 f(z_L) P_n(z_L) dz'$$
(5.48)

where the first few Legendre polynomials of interest to us are given explicitly as shown in equation (5.49). Figure 5.14 shows plots of these functions for $h_v = 10$ m. The first represents a simple uniform distribution, while the second includes linear variations, then quadratic and so on, with the higher-order



Fig. 5.14 The Legendre polynomials from zeroth to sixth order

functions offering ever-higher resolution of functional variation. In this way any function can be represented over the interval from z = 0 to $z = h_v$ by the 'spectrum' of real parameters a_n .

$$P_{0}(z) = 1$$

$$P_{1}(z) = z$$

$$P_{2}(z) = \frac{1}{2} (3z^{2} - 1)$$

$$P_{3}(z) = \frac{1}{2} (5z^{3} - 3z)$$

$$P_{4}(z) = \frac{1}{8} (35z^{4} - 30z^{2} + 3)$$

$$P_{5}(z) = \frac{1}{8} (63z^{5} - 70z^{3} + 15z)$$

$$P_{6}(z) = \frac{1}{16} (231z^{6} - 315z^{4} + 105z^{2} - 5)$$
(5.49)

The numerator and denominator of the general expression for coherence can now be written as shown in equation (5.50):

$$\int_{0}^{h_{v}} f(z')e^{i\beta_{z}z'}dz' = \frac{h_{v}}{2}e^{i\frac{\beta_{z}h_{v}}{2}}\int_{-1}^{1} (1+f(z_{L}))e^{i\frac{\beta_{z}h_{v}}{2}z_{L}}dz_{L}$$

$$\int_{0}^{h_{v}} f(z)dz = \frac{h_{v}}{2}\int_{-1}^{1} (1+f(z_{L}))dz_{L}$$
(5.50)

from which it follows that the coherence can be written as shown in equation (5.51):

$$\tilde{\gamma} = e^{i\beta_{z}z_{0}}e^{i\frac{\beta_{z}h_{v}}{2}}\frac{\int_{-1}^{1}(1+f(z_{L}))e^{i\frac{\beta_{z}h_{v}}{2}z_{L}}dz_{L}}{\int_{-1}^{1}(1+f(z_{L}))dz_{L}}$$

$$= e^{i\beta_{z}z_{0}}e^{i\beta_{v}}\frac{\int_{-1}^{1}\left(1+\sum_{n}a_{n}P_{n}(z_{L})\right)e^{i\beta_{v}z_{L}}dz_{L}}{\int_{-1}^{1}\left(1+\sum_{n}a_{n}P_{n}(z_{L})\right)dz_{L}}$$
(5.51)

By expanding the series and collecting terms, this equation can be rewritten in simplified form, as shown in equation (5.52):

$$\tilde{\gamma} = e^{i\beta_{z}z_{0}}e^{i\beta_{v}}\frac{(1+a_{0})\int_{-1}^{1}e^{i\beta_{v}z_{L}}dz_{L} + a_{1}\int_{-1}^{1}P_{1}(z_{L})e^{i\beta_{v}z_{L}}dz_{L} + a_{2}\int_{-1}^{1}P_{2}(z_{L})e^{i\beta_{v}z_{L}}dz_{L} + \cdots}{(1+a_{0})\int_{-1}^{1}dz_{L} + a_{1}\int_{-1}^{1}P_{1}(z_{L})dz_{L} + a_{2}\int_{-1}^{1}P_{2}(z_{L})dz_{L} + \cdots}$$
$$= e^{i\beta_{z}z_{0}}e^{i\beta_{v}}\frac{(1+a_{0})f_{0} + a_{1}f_{1} + a_{2}f_{2} + \dots a_{n}f_{n}}{(1+a_{0})}$$
$$= e^{i\beta_{z}z_{0}}e^{i\beta_{v}}(f_{0} + a_{10}f_{1} + a_{20}f_{2} + \cdots) \quad a_{i0} = \frac{a_{i}}{1+a_{0}}$$
(5.52)

Note that evaluation of the denominator is simplified by using the orthogonality of the Legendre polynomials. Evaluation of the numerator involves determination of the functions f_n , which are straightforward integrals employing repeated use of the following identity:

$$\int z^{n} e^{\beta z} dz = \frac{e^{\beta z}}{\beta} \left(z^{n} - \frac{nz^{n-1}}{\beta} + \frac{n(n-1)z^{n-2}}{\beta^{2}} \cdots \frac{(-1)^{n}n!}{\beta^{n}} \right)$$
(5.53)

As an example, equation (5.54) shows detailed calculation of the first two terms in the series. The first, corresponding to the zeroth-order Legendre polynomial, just yields a SINC function, while the first order linear polynomial gives a slightly more complicated function.

$$f_{0} = \frac{1}{2} \int_{-1}^{1} e^{i\beta_{v}z} dz = \left[\frac{e^{i\beta_{v}z}}{i\beta_{v}}\right]_{-1}^{1} = \frac{1}{i2\beta_{v}} \left(e^{i\beta_{v}z} - e^{-i\beta_{v}z}\right) = \frac{\sin\beta_{v}}{\beta_{v}}$$

$$f_{1} = \frac{1}{2} \int_{-1}^{1} z e^{i\beta_{v}z} dz = \left[\frac{e^{i\beta_{v}z}}{i\beta_{v}} \left(z - \frac{1}{i\beta_{v}}\right)\right]_{-1}^{1}$$

$$= \frac{1}{i\beta_{v}} \left(e^{i\beta_{v}z} + e^{-i\beta_{v}z}\right) - \frac{1}{(i\beta_{v})^{2}} \left(e^{i\beta_{v}z} - e^{-i\beta_{v}z}\right)$$

$$= i \left(\frac{\sin\beta_{v}}{\beta_{v}^{2}} - \frac{\cos\beta_{v}}{\beta_{v}}\right)$$
(5.54)

For reference we give the explicit form of all these functions up to sixth order in equation (5.55).

$$f_{o} = \frac{\sin \beta_{v}}{\beta_{v}}$$

$$f_{1} = i \left(\frac{\sin \beta_{v}}{\beta_{v}^{2}} - \frac{\cos \beta_{v}}{\beta_{v}} \right)$$

$$f_{2} = \frac{3 \cos \beta_{v}}{\beta_{v}^{2}} - \left(\frac{6 - 3\beta_{v}^{2}}{2\beta_{v}^{3}} + \frac{1}{2\beta_{v}} \right) \sin \beta_{v}$$

$$f_{3} = i \left(\left(\frac{30 - 5\beta_{v}^{2}}{2v\beta_{v}^{3}} + \frac{3}{2\beta_{v}} \right) \cos \beta_{v} - \left(\frac{30 - 15\beta_{v}^{2}}{2\beta_{v}^{4}} + \frac{3}{2\beta_{v}^{2}} \right) \sin \beta_{v} \right)$$

$$f_{4} = \left(\frac{35(\beta_{v}^{2} - 6)}{2\beta_{v}^{4}} - \frac{15}{2\beta_{v}^{2}} \right) \cos \beta_{v}$$

$$+ \left(\frac{35(\beta_{v}^{4} - 12\beta_{v}^{2} + 24)}{8\beta_{v}^{5}} + \frac{30(2 - \beta_{v}^{2})}{8\beta_{v}^{3}} + \frac{3}{8\beta_{v}} \right) \sin \beta_{v}$$

$$f_{5} = i \left(\frac{-2\beta_{v}^{4} + 210\beta_{v}^{2} - 1890}{\beta_{v}^{5}} \cos \beta_{v} + \frac{30\beta_{v}^{4} - 840\beta_{v}^{2} + 1890}{\beta_{v}^{6}} \sin \beta_{v} \right)$$

$$f_{6} = \left(\frac{42\beta_{v}^{4} - 2520\beta_{v}^{2} + 20790)}{\beta_{v}^{6}} \right) \cos \beta_{v}$$

$$+ \left(\frac{2\beta_{v}^{6} - 420\beta_{v}^{4} + 9450\beta_{v}^{2} - 20790)}{\beta_{v}^{7}} \right) \sin \beta_{v}$$
(5.55)

We note the following important points:

- 1. The even index functions are real while the odd are purely imaginary. We note also that the unknown coefficients a_n are all real.
- 2. The functions vary only with the single parameter β_v , which itself is defined from the product of two parameters: height h_v , and the interferometric wavenumber β_z .

Graphs of these functions are shown in Figure 5.15. We see that the first is a 'SINC' relation between coherence and increasing height-baseline product. This is the expected functional relationship for scattering by a uniform layer. However, we see that as the height-baseline product increases so the other functions become more important. We can conclude, therefore, that the interferometric coherence is sensitive to changes in the structure function f(z). There are two special cases of structure function of particular importance due to their widespread use in the literature. We now turn to consider these in more detail.

5.2.4.1 Special case 1: the uniform profile

If we assume f(z) = 1—a constant structure function—then all the higher order Legendre coefficients are zero, and the coherence becomes a function only of



Fig. 5.15 Coherence basis functions for Legendre expansion

height, given by a complex SINC function, as shown in equation (5.56):

$$\tilde{\gamma} = e^{i\beta_z z_0} e^{i\beta_z \frac{h_v}{2}} \frac{\sin\left(\frac{\beta_z h_v}{2}\right)}{\frac{\beta_z h_v}{2}}$$
(5.56)

There are two important features of this model. Firstly it shows that volume scattering provides a phase offset, given in this case by half the volume height. Hence in the presence of volume scattering the interferometric phase no longer represents the true surface position but is offset by a bias. For vegetated terrain this is called vegetation bias, and provides an error source in the use of radar interferometry for true surface topography mapping. We see that the only way to minimize this effect is to employ small baselines so that the product $\beta_z h_v$ remains small. However, this reduces the sensitivity of the interferometer and is difficult to sustain over forested terrain, where h_v can reach up to 50 m or more. Note that we cannot simply use the phase of the interferogram to estimate volume height h_v , since the total phase involves addition of an unknown phase shift due to the lower bound of the volume (z_0) . Only if we can provide an estimate of this lower 8 how to provide such an estimate.

The second key feature of the SINC model is that the coherence amplitude falls with increasing height and hence the phase variance increases with h_v . Note that there is no effect of the lower bounding surface on coherence amplitude (assuming range spectral filtering has been employed), and in principle we can therefore use an estimate of measured coherence amplitude to estimate height (for a known baseline). In particular, for short baselines we can expand the SINC function in a series and obtain a useful direct height estimate from coherence, as shown in equation (5.57):

$$x \ll 1 \Rightarrow \frac{\sin x}{x} \approx 1 - \frac{x^2}{6} \Rightarrow h_v \approx \sqrt{\frac{24(1 - |\tilde{\gamma}|)}{\beta_z^2}}$$
 (5.57)

However, as we shall see in the next section, this approach is sensitive to variations in the actual structure function of the volume. The SINC model is really valid only for very small height-baseline products, and for moderate baselines higher-order terms in the Legendre expansion of f(z) can no longer be ignored. In fact, as we shall see in Chapter 8, we can turn this idea around and design offset baselines to enhance the higher-order terms and thus enable parameter estimation for the layer.

Nonetheless, this SINC model is commonly used, especially by radar system designers, who wish only to assess the relative importance of volume decorrelation in the overall coherence budget for an interferometer. Finally, we note that this model contains no polarisation dependence at all. The volume decorrelation and phase bias of the SINC model are functions only of the height h_v . We shall see, however, that the higher-order terms of the Legendre expansion are sensitive to changes in wave polarisation, and this will suggest the development of polarimetric interferometry for parameter estimation. First, however, we turn to consider a second important special case: the exponential profile.

5.2.4.2 Special case 2: the exponential profile

A second important structure function is the exponential—widely used to model the physical effects of wave propagation through a volume scattering layer (Treuhaft, 1996, 2000a; Papathanassiou, 2001). This is in accordance with the water cloud model described in Section 3.5.1. According to this idea, contributions from the top of the volume are weighted more strongly in the coherence calculation than those deeper into the volume, as the latter experience a smaller incident signal due to wave extinction, combining the physical effects of wave attenuation due to absorption of energy by the volume and scattering loss due to the presence of particles. The combined effect of these two processes can be represented by a one-way *power* loss extinction coefficient σ_e with natural units of m^{-1} , but often expressed in engineering units of decibels per meter (dB/m). Note that two systems of units can be related using equation (5.58) (compare this with equation (3.11), for amplitude extinction). In addition we note that in radar applications there is a two-way propagation channel, and so the signal is attenuated both on the way in and out of the volume (see Figure 5.16). Hence the total extinction is $2\sigma_{e}$. Finally, we must also account for the increased attenuation path length through the medium when illuminated at an angle of incidence θ_0 , as shown in Figure 5.16.

$$\sigma_e^{dB} = \frac{10\sigma_e}{\ln(10)} \approx 4.34\sigma_e \Rightarrow \sigma_e \approx 0.23\sigma_e^{dB}$$
(5.58)

Rather than expand the exponential function in a Legendre series, it is easier in this case to explicitly evaluate the coherence integrals, as shown in



equation (5.59):

$$\hat{\gamma} = e^{i\beta_{z}z_{0}} \frac{m_{v} \int_{0}^{h_{v}} e^{\frac{2\sigma_{e}z_{o}}{\cos\theta_{o}}} e^{i\beta_{z}z} dz}{m_{v} \int_{0}^{h_{v}} e^{\frac{2\sigma_{e}z}{\cos\theta_{o}}} dz}
= \frac{2\sigma_{e} e^{i\beta_{z}z_{0}}}{\cos\theta_{o}(e^{2\sigma_{e}h_{v}/\cos\theta_{o}} - 1)} \int_{0}^{h_{v}} e^{i\beta_{z}z} e^{\frac{2\sigma_{e}z}{\cos\theta_{o}}} dz$$

$$= f(h_{v}, \sigma_{e}) = e^{i\beta_{z}z_{0}} \frac{p_{1}(e^{p_{2}h_{v}} - 1)}{p_{2}(e^{p_{1}h_{v}} - 1)} \qquad \begin{cases} p_{1} = \frac{2\sigma_{e}}{\cos\theta_{o}} \\ p_{2} = \frac{2\sigma_{e}}{\cos\theta_{o}} + i\beta_{z} \end{cases}$$
(5.59)

This example illustrates the important new idea that the coherence in general depends not only on the volume depth h_v but also on the shape of the vertical structure function. The exponential model essentially allows a one-parameter model for variation of structure (via σ_e). High extinction implies an effective scattering layer at the top of the volume, such as a high-elevated forest canopy, for example. Figure 5.17 shows an example of how the coherence varies for an exponential profile with varying σ_e and depth h_v . We have selected a baseline corresponding to $\beta_z = 0.1567$ (which corresponds to a zero of the SINC model at 40 m), and considered a 45-degree angle of incidence. Note that for zero extinction we again obtain, as a special case, the SINC model.

However, as extinction increases so the coherence increases for a given height. This arises physically as the effective scattering volume is being squeezed into a smaller and smaller region close to the top of the volume as extinction is increased. This can be confirmed by plotting the phase of the coherence. We first define the fractional phase centre height P_c from the interferometric phase ϕ as $P_c = \frac{\phi}{h_c \beta_c}$. Figure 5.18 shows how P_c varies with extinction



Fig. 5.17 Volume decorrelation versus height for various extinctions



Fig. 5.18 Phase centre height versus height for various extinctions



and height. Note that for the special case of zero extinction we obtain a phase centre halfway up the layer as expected in the SINC model. However, as the extinction increases we see that the phase centre moves towards the top of the layer, approaching $P_c = 1$ in the limit of infinite extinction.

We have seen from Figures 5.17 and 5.18 that the coherence amplitude and phase variations are linked. Indeed, it is instructive to visualize both at the same time by employing the coherence diagram representation (see Appendix 3). Figure 5.19 shows how the complex coherence varies inside the unit circle in the complex coherence plane for three extinction values and layer depths varying from 0 to 40 m (using the same parameters used in Figures 5.17 and 5.18). Here we see that the SINC model spirals quickly to the origin—to zero

coherence—while the high-extinction cases show gentler spirals with more rapid phase variation around the unit circle.

5.2.5 Summary: coherence decomposition

We have seen in this chapter that the interferometric coherence may be decomposed into a product of terms, the most important of which are shown in equation (5.60), where we define the following important components:

$$\tilde{\gamma} = e^{i\phi_s} \gamma_{SNR} \gamma_t \gamma_{proc} \gamma_s \tilde{\gamma}_v \tag{5.60}$$

 γ_{SNR} Decorrelation due to additive noise in the signals.

- γ_t Temporal decorrelation due to motion of scatterers between passes in repeat-pass interferometry.
- γ_{proc} Loss of coherence due to processing errors associated, for example, with image misregistration in radar imaging.
 - γ_s Baseline or surface decorrelation. This depends on the nature of the surface scattering (point scatterers or random surface scattering), but can always be removed (set equal to 1) by employing range spectral filtering.
 - $\tilde{\gamma}_v$ Volume decorrelation. This is a complex coherence, in that unlike the other terms it distorts both the mean and standard deviation of the interferometric phase.

We have developed a general method for predicting the volume coherence for a given structure function using a generalized Fourier–Legendre expansion, and considered in detail the important special cases of a uniform and exponential profile.

In particular we have seen that interferometric coherence can be controlled through baseline selection, even for scattering from random media. This gives us the 'entropy control' missed with polarimetry alone. The next step is to incorporate polarisation effects into radar interferometry. In Chapter 6 we show how to do this in a formal mathematical way before exploring some of the physical models used in Chapter 7.

6

Polarimetric interferometry

In this chapter we formally combine the topics of polarimetry and interferometry. Our purpose is to establish a general framework for describing the formation and analysis of interferograms for arbitrary choice of transmit and receive wave polarisations (Papathanassiou, 1997). This will lead us to study the variation of interferometric coherence with polarisation, and ultimately to develop methods for coherence optimization (Cloude, 1997b), for investigating the dynamic range of interferometric coherence variation with polarisation. This will then lead us, in Chapter 7, to apply the optimization procedures to surface and volume scattering scenarios in the same way as for polarimetry alone in Chapters 3 and 4.

6.1 Vector formulation of radar interferometry

To generate a vector interferogram we require two key ideas (Cloude, 1997b, 1998). The first is that an interferogram is always formed between two complex scalars, representing the amplitude and phase of scattered fields at ends 1 and 2 of a spatial or temporal baseline. We therefore need some general way to project the vector polarisation matrix data onto a complex scalar quantity. The standard way to do this is through a Hermitian inner product of vectors $s = \underline{x}^{*T}$ *y*. This has the advantage that it directly yields a scalar phase related to the differences between \underline{x} and *y*.

The second key idea is that we can always select an arbitrary polarimetric scattering mechanism using the <u>w</u> vector formulation, introduced in Chapter 2 and shown again for N = 1, 2, 3 and 4-dimensional scattering in equation (6.1). Conventional single channel 'scalar' interferometry then makes use of <u>w</u>(1), dual and compact polarimetry <u>w</u>(2), full backscatter polarimetry <u>w</u>(3), and bistatic polarimetry <u>w</u>(4).

$$\underline{w}(1) = e^{i\phi} \quad \underline{w}(2) = \begin{pmatrix} \cos \alpha & e^{i\phi_1} \\ \sin \alpha & e^{i\phi_2} \end{pmatrix}$$
$$\underline{w}(3) = \begin{pmatrix} \cos \alpha & e^{i\phi_1} \\ \sin \alpha \cos \psi & e^{i\phi_2} \\ \sin \alpha \sin \psi & e^{i\phi_3} \end{pmatrix} \quad \underline{w}(4) = \begin{pmatrix} \cos \alpha & e^{i\phi_1} \\ \sin \alpha \cos \psi & e^{i\phi_2} \\ \sin \alpha \sin \psi & \cos \gamma & e^{i\phi_3} \\ \sin \alpha \sin \psi & \sin \gamma & e^{i\phi_4} \end{pmatrix}$$
(6.1)

Combining these two ideas leads us to the following general procedure for generating a vector interferogram. We first project the complex scattering vectors \underline{k}_1 and \underline{k}_2 , measured at ends 1 and 2 of the baseline, onto the conjugate of

the desired polarimetric scattering mechanisms \underline{w}_1 and \underline{w}_2 (which importantly may be different polarisations at either end of the baseline). These projections provide two complex scalars s_1 and s_2 , representing the complex scattering components that can then be combined into an interferogram, from which we can estimate the corresponding phase using a standard Hermitian inner product for complex vectors, as shown in equation (6.2):

$$s_1 = \underline{w}_1^{*T} \underline{k}_1$$

$$s_2 = \underline{w}_2^{*T} \underline{k}_2$$

$$\Rightarrow \phi = \arg(s_1 s_2^*) = \arg(\underline{w}_1^{*T} \underline{k}_1 \underline{k}_2^{*T} \underline{w}_2)$$
(6.2)

This expression is quite general, applying to N = 1, 2, 3 or 4 depolarisation problems. The most common form used in radar is the N = 3 case (backscatter with reciprocity), explicitly shown for reference in the Pauli base in equation (6.3):

$$s_{1} = w_{1,1}^{*} \frac{(s_{hh}^{1} + s_{vv}^{1})}{\sqrt{2}} + w_{1,2}^{*} \frac{(s_{hh}^{1} - s_{vv}^{1})}{\sqrt{2}} + w_{1,3}^{*} \sqrt{2} s_{hv}^{1} = \underline{w}_{1}^{*T} \underline{k}_{1}$$

$$s_{2} = w_{2,1}^{*} \frac{(s_{hh}^{2} + s_{vv}^{2})}{\sqrt{2}} + w_{2,2}^{*} \frac{(s_{hh}^{2} - s_{vv}^{2})}{\sqrt{2}} + w_{2,3}^{*} \sqrt{2} s_{hv}^{2} = \underline{w}_{2}^{*T} \underline{k}_{2}$$

$$\Rightarrow \phi = \arg(s_{1}s_{2}^{*}) = \arg(\underline{w}_{1}^{*T} \underline{k}_{1} \underline{k}_{2}^{*T} \underline{w}_{2})$$
(6.3)

Before proceeding, one important required piece of housekeeping is that we do not want the phase of the interferogram to depend on the arbitrary phase difference between the complex vectors \underline{w}_1 and \underline{w}_2 , and we therefore enforce the additional normalization constraint shown in equation (6.4):

$$\phi_w = \arg\left(\underline{w}_1^{*T}\underline{w}_2\right) = 0 \tag{6.4}$$

This is automatically satisfied if we choose $\underline{w}_1 = \underline{w}_2$, but in the general case must be explicitly enforced by modifying \underline{w}_2 , as shown in equation (6.5):

$$\underline{w}_2 \to e^{-i \arg(\underline{w}_1^{*T} \underline{w}_2)} \underline{w}_2 \tag{6.5}$$

In this way we can include polarimetry with interferometry in a consistent, complete and logical manner for any dimension of depolarisation. Importantly, this same approach can then be combined with averaging to predict the coherence of the interferogram for polarisations \underline{w}_1 and \underline{w}_2 , as shown in equation (6.6):

$$\begin{cases} s_1 = \underline{w}_1^{*T} \cdot \underline{k}_1 \\ s_2 = \underline{w}_2^{*T} \cdot \underline{k}_2 \end{cases} \Rightarrow \tilde{\gamma} \left(\underline{w}_1, \underline{w}_2 \right) = \frac{E(s_1 s_2^*)}{\sqrt{E(s_1 s_1^*)} \cdot \sqrt{E(s_2 s_2^*)}} \quad 0 \le |\tilde{\gamma}| \le 1 \quad (6.6)$$

6.1.1 Generalized coherency matrix formulation

We can reformulate this procedure for complex coherence estimation using matrices, as shown in equation (6.7). The advantage of doing this is that we can then easily extend the idea to multiple baselines and also provide a formal link with our ideas about wave depolarisation and coherence in different dimensions. The basic idea is to stack the coherent scattering vectors for each end of the baseline, \underline{k}_1 and \underline{k}_2 , into a single column vector. The coherency matrix is then formed from the average product of this vector with its conjugate

transpose, called $[\Lambda_2]$, where the subscript 2 now refers to the number of spatial positions used.

$$[\Lambda_2] = \left\langle \begin{bmatrix} \underline{k}_1 \\ \underline{k}_2 \end{bmatrix} \cdot \begin{bmatrix} \underline{k}_1^{*T} & \underline{k}_2^{*T} \end{bmatrix} \right\rangle = \begin{bmatrix} T_{11} & \Omega_{12} \\ \Omega_{12}^{*T} & T_{22} \end{bmatrix}$$

$$\Rightarrow \quad \tilde{\gamma}(\underline{w}_1, \underline{w}_2) = \frac{\underline{w}_1^{*T} \Omega_{12} \underline{w}_2}{\sqrt{\underline{w}_1^{*T} T_{11} \underline{w}_1} \cdot \sqrt{\underline{w}_2^{*T} T_{22} \underline{w}_2}}$$

$$(6.7)$$

For N-dimensional depolarisation problems $[\Lambda_2]$ is a 2N × 2N Hermitian matrix. In the general multibaseline case, where M spatial positions are available, the matrix $[\Lambda_M]$ becomes MN × MN in size.

We can make one further structural observation about the general Hermitian matrix $[\Lambda_M]$. It is always composed of N × N sub-matrices, as shown in equation (6.8), where the M diagonal blocks T_{ii} represent the polarimetric information at each of the M spatial positions. The information in these matrices can be interpreted using any of the depolarisation techniques (such as entropy/alpha) discussed in Chapters 2 and 4.

$$[\Lambda_1] = [T] \rightarrow [\Lambda_2] = \begin{bmatrix} T_1 & \Omega_{12} \\ \Omega_{12}^* & T_2 \end{bmatrix} \rightarrow [\Lambda_M] = \begin{bmatrix} T_{11} & \Omega_{12} & \dots & \Omega_{1M} \\ \Omega_{12}^* & T_{22} & \dots & \Omega_{2M} \\ \vdots & \vdots & \ddots & \vdots \\ \Omega_{1M}^* & \Omega_{2M}^* & \dots & T_{MM} \end{bmatrix}$$
(6.8)

Here our interest centres more on the new N × N complex matrices $[\Omega_{ij}]$, which contain information related to the variation of *interferometric* phase with polarisation. These matrices are neither Hermitian nor unitary, and hence have a general 3 × 3 complex structure. We can see that these block elements play an important and separable role in determining the coherence, as shown at right in equation (6.7). Under a unitary change of base of the scattering vector \underline{k} by an N × N unitary matrix, $[\Lambda_2]$ then transforms as shown in equation (6.9):

$$\underline{k}' = \begin{bmatrix} U_N \end{bmatrix} \underline{k} \Rightarrow \begin{bmatrix} T'_{11} & \Omega'_{12} \\ \Omega_{12}^{*T} & T'_{22} \end{bmatrix} = \begin{bmatrix} U_N & 0 \\ 0 & U_N \end{bmatrix} \begin{bmatrix} T_{11} & \Omega_{12} \\ \Omega_{12}^{*T} & T_{22} \end{bmatrix} \begin{bmatrix} U_N^{*T} & 0 \\ 0 & U_N^{*T} \end{bmatrix}$$
(6.9)

From here we can then start by considering single baseline polarimetric interferometry (SBPI) [Λ_2], which involves measurements at only two separated spatial/temporal positions. Here the T_{ii} matrices can still have polarimetric dimension N = 1, 2, 3 or 4, but the matrix Ω_{12} contains additional information about the variation of interferometric coherence. This procedure can then be easily extended to multiple baselines (and frequencies), as shown in equation (6.8) (Ferro-Famil, 2001, 2008).

Considering the special but important case of radar backscatter N = 3, there are several special cases of unitary change of base U_3 to be distinguished. Equation (6.9) is expressed in the linear Pauli basis (see equation (6.3)). If we wish to convert to the standard linear lexicographic base of HH, $\sqrt{2}$ HV and VV, to predict the interferometric coherence in these channels, then we can employ the following unitary matrix in equation (6.9) (see equation (2.45)):

$$U_N = U_{LP3} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 & 0\\ 0 & 0 & \sqrt{2}\\ 1 & -1 & 0 \end{bmatrix}$$
(6.10)

We can then focus attention on a general change of wave base states: movement of a reference point *P* over the surface of the Poincaré sphere. If the spherical triangle coordinates of *P* are α_w and δ_w (see Figure 1.12), then the general unitary matrix for use in equation (6.9) takes the form shown in equation (6.11):

$$P = \begin{bmatrix} \cos \alpha_w \\ \sin \alpha_w e^{i\delta_w} \end{bmatrix} \Rightarrow [U_3] = [U_3^L][U_{LP3}]$$

$$\Rightarrow [U_3^L] = \begin{bmatrix} \cos^2 \alpha_w & -\sqrt{2}\cos \alpha_w \sin \alpha_w e^{-i\delta_w} & \sin^2 \alpha_w e^{-i2\delta_w} \\ \sqrt{2}\cos \alpha_w \sin \alpha_w e^{i\delta_w} & \cos^2 \alpha_w - \sin^2 \alpha_w & -\sqrt{2}\cos \alpha_w \sin \alpha_w e^{-i\delta_w} \\ \sin^2 \alpha_w e^{i2\delta_w} & \sqrt{2}\cos \alpha_w \sin \alpha_w e^{i\delta_w} & \cos^2 \alpha_w \end{bmatrix}$$
(6.11)

For example, if we want to convert from the linear H,V basis to left and right circular L,R so as to obtain the matrix in the basis LL, $\sqrt{2}$ LR,RR, then we would set $\alpha_w = \pi/4$, $\delta_w = \pi/2$ in equation (6.11) and obtain the composite change of basis matrix shown in equation (6.12). This can then be used in equation (6.9) to express all matrices in the circular basis.

$$\begin{bmatrix} U_3^{circ} \end{bmatrix} = \frac{1}{2} \begin{bmatrix} 1 & \sqrt{2}i & -1 \\ \sqrt{2}i & 0 & \sqrt{2}i \\ -1 & \sqrt{2}i & 1 \end{bmatrix} \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 & 0 \\ 0 & 0 & \sqrt{2} \\ 1 & -1 & 0 \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} 0 & 1 & i \\ \sqrt{2}i & 0 & 0 \\ 0 & -1 & i \end{bmatrix}$$
(6.12)

Note, however, that equation (6.11) does not represent the most general unitary transformation. It has only two free parameters, while the general 3×3 unitary matrix has nine (see Appendix 2). These extra degrees of freedom are generated by combining triplets of orthogonal scattering mechanisms, as used in the eigenvector decomposition of [T], for example (see Section 4.22). Starting with an arbitrary mechanism, we have five degrees of freedom; the second then must be orthogonal to the first, and so has 5 - 2 = 3 parameters. The third must then be orthogonal to the first two, and so has 5 - 4 = 1 parameter, producing nine in total. If these are then combined into a special unitary matrix (det(U_3) = 1) this reduces to eight parameters. The set of general unitary transformations is then governed by the eight Gell–Mann matrices, as shown in equation (6.13) (see Appendix 2) (Cloude, 1995b; Ferro-Famil, 2000).

$$[U_3] = \begin{bmatrix} \underline{w}_1 & \underline{w}_2 & \underline{w}_3 \end{bmatrix} \stackrel{\det(U_3)=1}{\longrightarrow} [U_3] = \exp(i\phi\underline{n}.\underline{G})$$
(6.13)

The key conclusion is that arbitrary unitary matrices in the change of base formulation—equation (6.9)—can be given a clear physical interpretation in terms of triplets of polarimetric scattering mechanisms. The change of wave polarisation base then forms only a subset of these matrices through equation

Polarisation selection/w vector	w ₁	w ₂	w ₃
НН	Ö2	Ö2	0
HV	0	0	1
VV	Ö2	-Ö2	0
HH+VV	1	0	0
HH+VV	0	1	0
2HV	0	0	1
LL	0	Ö2	Ö2i
LR	1	0	0
RR	0	-Ö2	Ö2i

Fig. 6.1 Example scattering mechanisms used for POLInSAR

(6.11). This will be important when we come to consider coherence optimization in Section 6.2, as we can then allow unconstrained search through all available parameters of the unitary matrix.

While these formal unitary transformations are useful for analytical manipulations, in practice we are very often concerned only with direct evaluation of the coherence (equation (6.7)) for different polarisations. In this case we can first estimate the matrices in a fixed basis (the Pauli basis of equation (6.3), for example), and then use diversity of \underline{w} to generate the different polarisations. The weight vectors \underline{w}_1 and \underline{w}_2 then define user-selected scattering mechanisms at ends 1 and 2 of the across-track baseline. Figure 6.1 shows some important examples of the weight vector $\underline{w} = (w_1, w_2, w_3)^T$ for coherence estimation in the commonly used linear, Pauli and circular bases. This table can be used with equation (6.7) to generate interferograms in different polarisation channels.

Consequently, in applications we first need to estimate the composite matrices of equation (6.8) from the radar data itself. Given L samples of MN dimensional scattering vectors \underline{u} , the estimate [Z] of [Λ] is then conveniently formed using a maximum likelihood (ML) estimator, as shown in equation (6.14):

$$[Z] = \frac{1}{L} \sum_{j=1}^{L} \underline{u}_j \underline{u}_j^{*T}$$
(6.14)

For finite L there will be errors in this estimate relating to higher-dimensional forms of coherence bias, as discussed in Appendix 3. To illustrate this, consider a numerical example from single baseline polarimetric radar interferometry (SBPI), when MN = 6, as shown in equation (6.15):

$$[\Lambda_2] = \begin{bmatrix} 2 & 0 & 0 & 1.8 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0.6e^{i\frac{\pi}{4}} & 0 \\ 0 & 0 & 1 & 0 & 0 & 0.4e^{i\frac{\pi}{2}} \\ 1.8 & 0 & 0 & 2 & 0 & 0 \\ 0 & 0.6e^{-i\frac{\pi}{4}} & 0 & 0 & 1 & 0 \\ 0 & 0 & 0.4e^{-i\frac{\pi}{2}} & 0 & 0 & 1 \end{bmatrix}$$
(6.15)

This matrix corresponds physically to scattering by a random dipole cloud with polarisation dependent complex interferometric coherences of magnitude 0.9, 0.6 and 0.4 respectively, and with separated interferometric phase centres



Fig. 6.2 Estimation of coherence triplet of equation (6.15) versus number of looks L

of 0, $\pi/4$ and $\pi/2$. If we now consider numerical estimation of these three coherence amplitudes as a function of number of looks L, using the Monte Carlo data simulation technique described in Appendix 3, we obtain the typical convergence shown in Figure 6.2. Each point in this graph (for a fixed L) is obtained as the mean of 256 realizations of the randomized estimation process. We see a small noise variation due to the finite sampling, but can still see the general behaviour expected of coherence estimation: namely, coherence bias for a small number of looks, which reduces as L increases. This bias also increases with decreasing coherence, so we see the 0.9 channel has little bias, and we obtain accurate estimates for a small number L > 5 looks. The low 0.4 channel, however, shows much slower convergence, and requires in excess of 25 looks for 'good' estimation.

This bias is due to finite sampling, which also has an impact on the estimation of depolarisation parameters of the cloud. For example, the scattering entropy of the cloud of dipoles is H = 0.946 (see Figure 3.29). We can estimate this entropy again as a function of L by isolating the T_{11} component of the estimated coherency matrix, performing an eigenvalue analysis and then calculating entropy. When we do this as a function of number of looks L, we obtain the estimates shown in Figure 6.3. Note that in this case the entropy estimate is underestimated for a small number of looks, and only slowly converges to the correct value (we obtain around 5% relative error for L > 20 looks). We note that this bias depends on the underlying entropy. For low entropy scattering (with one dominant eigenvalue) the convergence accelerates, and the multilooking requirements are much reduced. From this point of view a dipole cloud represents an extreme case of depolarisation, and hence represents an upper bound on the bias issues for single scattering problems.

Nonetheless, such numerical biases have to be considered when dealing with practical questions of the variation of coherence with polarisation, as clearly the apparent dynamic range will always depend on L and we must ensure that


Fig. 6.3 Estimated scattering entropy versus number of looks for dipole cloud

L is sufficiently large to minimize any numerical bias. We now turn to consider the issue of quantifying the range of coherence variation with polarisation by employing systematic optimization techniques based on Lagrange multipliers.

6.2 Coherence optimization

A fundamental question of importance in polarimetric interferometry is to determine the maximum interferometric coherence change with polarisation. If it changes only slightly, then polarimetry plays only a weak role. On the other hand, if the coherence varies strongly with polarisation then this indicates important changes in the relative positions of scattering mechanisms, which we can then exploit for parameter estimation. A quantitative approach to this estimation can be made based on the mathematics of optimization theory to which we now turn (Cloude, 1997b, 1998; Tabb, 2001, 2002a, 2002b; Pascual, 2002; Colin, 2005, 2006; Neumann, 2008).

We first investigate this question by using a formal Lagrange multiplier optimization process as follows. Our starting point is the general expression for complex coherence, conveniently written in terms of sub-matrices, as shown in equation (6.16):.

$$\tilde{\gamma}\left(\underline{w}_{1},\underline{w}_{2}\right) = \frac{\underline{w}_{1}^{*T}\Omega_{12}\underline{w}_{2}}{\sqrt{\underline{w}_{1}^{*T}T_{11}\underline{w}_{1}}\cdot\underline{w}_{2}^{*T}T_{22}\underline{w}_{2}}} \to \max_{\underline{w}_{1}\underline{w}_{2}}|\tilde{\gamma}|$$
(6.16)

This represents the coherence obtained when polarisation \underline{w}_1 is used at the first and \underline{w}_2 at the second end of the baseline. In general, therefore, it combines both interferometric and polarimetric contributions to coherence. Our objective is to find the extreme values of the magnitude of this function. There is a slight complication in that equation (6.16) is a complex function, and so we must decide whether to maximize the absolute value, the phase, or real and imaginary parts. These various choices open up different forms of optimization as we now consider.

6.2.1 Unconstrained optimization

We start by determining which scattering mechanisms \underline{w}_1 and \underline{w}_2 maximize the *magnitude* of the interferometric coherence (Cloude, 1997b). To answer this, we set up a (complex) Lagrangian function *L*, as shown in equation (6.17). This function comprizes the numerator of the coherence constrained by two Lagrange parameters λ_1 and λ_2 , which permit variation of the numerator while keeping the denominator constant. In this way we can find the extreme values of the magnitude *LL** by setting the complex partial derivatives of *L* and *L** to zero, as shown in equation (6.17):

$$L = \underline{w}_1^{*T} \Omega_{12} \underline{w}_2 + \lambda_1 \left(\underline{w}_1^{*T} T_{11} \underline{w}_1 - 1 \right) + \lambda_2 \left(\underline{w}_2^{*T} T_{22} \underline{w}_2 - 1 \right)$$

$$\Rightarrow \begin{cases} \frac{\partial L}{\partial \underline{w}_1^{*T}} = \Omega_{12} \underline{w}_2 + \lambda_1 T_{11} \underline{w}_1 = 0 \\ \frac{\partial L^*}{\partial w_2^{*T}} = \Omega_{12}^{*T} \underline{w}_1 + \lambda_2^* T_{22} \underline{w}_2 = 0 \end{cases}$$
(6.17)

This yields a set of coupled equations for the unknown vectors \underline{w}_1 and \underline{w}_2 and the Lagrange multipliers λ_1 and λ_2 . There are now two important options for solution of these equations. In the most general case we allow \underline{w}_1 and \underline{w}_2 to be different and so allow full polarisation diversity. In this case we can find a solution to the coupled equations as a pair of eigenvalue problems, as shown in equation (6.18):

$$\stackrel{\underline{w}_{1}\neq\underline{w}_{2}}{\longrightarrow} \begin{cases} T_{22}^{-1}\Omega_{12}^{*T}T_{11}^{-1}\Omega_{12}\underline{w}_{2} = \lambda_{1}\lambda_{2}^{*}\underline{w}_{2} \\ T_{11}^{-1}\Omega_{12}T_{22}^{-1}\Omega_{12}^{*T}\underline{w}_{1} = \lambda_{1}\lambda_{2}^{*}\underline{w}_{1} \end{cases} \qquad \lambda_{1} = \lambda_{2} = \tilde{\gamma}_{opt} \Rightarrow K_{2}\underline{w}_{2} = v\underline{w}_{2} = \left|\gamma_{opt}\right|^{2}\underline{w}_{2} \\ K_{2}\underline{w}_{2} = v\underline{w}_{2} = \left|\gamma_{opt}\right|^{2}\underline{w}_{2} \end{cases}$$

$$(6.18)$$

This shows that the optimum scattering mechanisms can be obtained from eigenvalue equations involving composite products of the elements of $[\Lambda_v]$. Furthermore, the two Lagrange multipliers are complex but equal. To show this we left multiply the top derivative equation in (6.17) by \underline{w}_1^{*T} and the lower by w_2^{*T} , and use the normalization condition on the Hermitian forms on the righthand side of L in equation (6.17) to show that $\lambda_1 = \lambda_2$. Note that for backscatter problems there are then three optimum values corresponding to the square moduli of the three eigenvalues of **K**. Hence the 3×3 matrices **K**₁ and **K**₂ have the same non-negative real eigenvalue spectra in the range $0 \le v_3 \le v_2 \le v_1 \le 1$, but different and non-orthogonal eigenvectors, as neither K_1 nor K_2 are generally Hermitian or unitary matrices. The maximum coherence is given by the square root of the largest eigenvalue v_1 with corresponding scattering mechanisms given by the eigenvectors. Note that the optimum complex coherence can be found by first calculating the eigenvectors w_1 and w_2 from equation (6.18), phase normalizing using equation (6.4), and then using them directly in equation (6.16). Alternatively we may calculate the optimum directly by solving the following generalized eigenvalue problem, obtained by a straightforward

rewriting of the derivative equations in (6.17) and using the fact that $\lambda_1 = \lambda_2$, as shown in equation (6.19):

$$\begin{bmatrix} 0 & \Omega_{12} \\ \Omega_{12}^{*T} & 0 \end{bmatrix} \begin{bmatrix} \underline{w}_1 \\ \underline{w}_2 \end{bmatrix} = \lambda \begin{bmatrix} T_{11} & 0 \\ 0 & T_{22} \end{bmatrix} \begin{bmatrix} \underline{w}_1 \\ \underline{w}_2 \end{bmatrix} \Rightarrow \tilde{\gamma}_{opt} = \lambda_{\max} e^{-i(\arg(\underline{w}_1^{*T} \underline{w}_2))}$$
(6.19)

where λ_{max} is the eigenvalue with maximum modulus. The advantage of this formulation is that it scales naturally to the multi-baseline case, as first shown in Neumann (2008). When M-tracks are available we must use the generalized coherency matrix $[\Lambda_M]$ shown in equation (6.8). In this case the unconstrained optimization problem can be formulated by generalising the Lagrangian to a sum of numerators with constrained denominators, leading to the generalization of equation (6.19), as shown in equation (6.20):

$$L = \sum_{i=1}^{M} \sum_{j=i+1}^{M} \underline{w}_{i}^{*T} \Omega_{ij} \underline{w}_{j} + \lambda \sum_{i=1}^{M} \left(\underline{w}_{i}^{*T} T_{ii} \underline{w}_{i} - 1 \right)$$

$$\Rightarrow \begin{bmatrix} 0 & \Omega_{12} & \dots & \Omega_{1M} \\ \Omega_{12}^{*T} & 0 & \dots & \Omega_{2M} \\ \vdots & \vdots & \ddots & \vdots \\ \Omega_{1M}^{*T} & \Omega_{2M}^{*T} & \dots & 0 \end{bmatrix} \begin{bmatrix} \underline{w}_{1} \\ \underline{w}_{2} \\ \vdots \\ \underline{w}_{M} \end{bmatrix} = \lambda \begin{bmatrix} T_{11} & 0 & \dots & 0 \\ 0 & T_{22} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & T_{MM} \end{bmatrix} \begin{bmatrix} \underline{w}_{1} \\ \underline{w}_{2} \\ \vdots \\ \underline{w}_{M} \end{bmatrix}$$

$$(6.20)$$

Note that here λ_{max} now corresponds to the weighted sum of optimized coherence moduli—a type of average across all the baselines.

6.2.1.1 SVD interpretation of unconstrained optimization

We can obtain a useful physical interpretation of this optimization process by reformulating it as a singular value decomposition (SVD) (see Appendix 1). The starting point for this is to realize that we can always pre-whiten the polarimetric scattering vectors; that is, we can transform them into a base with the identity as a coherency matrix, corresponding to 'white' noise. This can be achieved using a transformation involving the square root of the actual polarimetric coherency matrix, which can best be evaluated in terms of its matrix of eigenvalues [D] and eigenvectors [U], as shown in equation (6.21). This represents a change of polarisation base given by the matrix [U] followed by a weighting of the channels by the reciprocal of the square root of eigenvalues.

$$\underline{k}^{n} = \sqrt{T^{-1}\underline{k}} = \sqrt{D^{-1}} [U] \underline{k} \Rightarrow \langle \underline{k}^{n} \cdot \underline{k}^{n*T} \rangle = I_{N} = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 1 \end{bmatrix}$$
$$\Rightarrow \Pi = \sqrt{T_{11}^{-1}} \Omega_{12} \sqrt{T_{22}^{-1}}$$
$$\Lambda_{2}]_{noise} = \begin{bmatrix} I_{3} & \Pi \\ \Pi^{*T} & I_{3} \end{bmatrix} \Rightarrow \begin{cases} \Pi^{*T} \Pi \underline{w}_{2} = \lambda_{1} \lambda_{2}^{*} \underline{w}_{2} \\ \Pi \cdot \Pi^{*T} \underline{w}_{1} = \lambda_{1} \lambda_{2}^{*} \underline{w}_{1} \end{cases}$$
(6.21)

[

More significant is the effect of this transformation on the polarimetric interferometry sub-matrix $[\Omega_{12}]$. The transformation does not generate noise in the interferogram, but yields a structured matrix Π as shown in the lower part of equation (6.21). The optimum states \underline{w}_1 and \underline{w}_2 are then given as the left and right singular vectors of the matrix Π as shown. They can be obtained from standard eigenvalue problems for the Hermitian matrices $\Pi \Pi^{*T} \Pi$ and $\Pi^{*T} \Pi$ respectively. Coherence optimization can therefore be considered a problem in singular value decomposition of the matrix Π —physically the result of pre-whitened noise interferometry between the polarimetric channels.

6.2.2 Constrained optimization

A second important form of coherence optimization first imposes the additional constraint that $\underline{w}_1 = \underline{w}_2$ —that the scattering mechanisms at either end of the baseline are equal (Tabb, 2001, 2002a; Colin, 2005, 2006). This is often supported by the physical argument that for small baselines the optimum scattering mechanisms, in the absence of temporal changes, should be equal. However, as we shall see, there are also good numerical as well as physical reasons for adopting this approach in many applications.

In the constrained case the general optimization equations of (6.17) simplify as shown in equation (6.22):

$$\frac{w_{1}=w_{2}}{\longrightarrow} \begin{cases} \Omega_{12}\underline{w} + \lambda_{1}T_{11}\underline{w} = 0\\ \Omega_{12}^{*T}\underline{w} + \lambda_{2}^{*}T_{22}\underline{w} = 0 \end{cases} \Rightarrow (T_{11} + T_{22})^{-1} \left(\Omega_{12} + \Omega_{12}^{*T}\right)\underline{w} = -(\lambda_{1} + \lambda_{2}^{*})\underline{w} \\ [T]^{-1}[\Omega_{H}]\underline{w} = \lambda \left(\phi\right)\underline{w} \begin{cases} [\Omega_{H}] = \frac{1}{2} \left(\Omega_{12}e^{i\phi} + \Omega_{12}^{*T}e^{-i\phi}\right)\\ [T] = \frac{1}{2} \left(T_{11} + T_{22}\right) \end{cases} \\ \frac{\max|\lambda(\phi)|}{\longrightarrow} \underline{w}_{opt} \Rightarrow \gamma_{opt} = \frac{\underline{w}_{opt}^{*T}[\Omega_{H}]\underline{w}_{opt}}{\underline{w}_{opt}^{*T}[T]\underline{w}_{opt}} \end{cases}$$
(6.22)

We see again an eigenvalue equation, but this time based on averages of the sub-matrices. However, one drawback of this approach is, as shown on the right-hand side of equation (6.22), that it maximizes only the real part of the eigenvalue and so represents a phase-sensitive optimization. Hence this finds only a local maximum, and to find the true global optima we need to introduce a free phase parameter $\exp(i\phi)$. By then repeating the optimization in equation (6.22) for different values of ϕ we can then obtain the global maxima. The general procedure for constrained optimization is then summarized in the lower portion of equation (6.22).

The optimization process is then formally equivalent to a mathematical property called the numerical radius of an N × N complex matrix [A] (Murnaghan, 1932; Li, 1994; He, 1997; Mengi, 2005). This itself is defined from the field of values F(A), as defined in equation (6.23):

$$F(A) = \left\{ \underline{w}^{*T}[A]\underline{w}, \underline{w} \in C^N, \left\| \underline{w} \right\| = 1 \right\}$$
(6.23)

The numerical radius is then the radius of the smallest circle that contains the field of values, as defined in equation (6.24):

$$r(A) = \max\{|z|, z \in F(A)\}$$
(6.24)

In our context we can formally relate the numerical radius to coherence optimization by first generating the constrained form of the Lagrangian function L_C , as shown in equation (6.25):

$$L_C = \underline{w}^{*T} \Omega_{12} \underline{w} + \lambda \left(\underline{w}^{*T} T \underline{w} - 1 \right)$$
(6.25)

This is almost in the form required, and needs only a pre-whitening transformation to remove the polarisation dependence of the constraint equation (the factor \mathbf{T}). Using the square root of \mathbf{T} as a basis transformation, we then obtain the following modified form:

$$\underline{w}_n = \sqrt{T^{-1}}\underline{w} \Rightarrow L_c = \underline{w}_n^{*T} T^{-\frac{1}{2}} \Omega_{12} T^{-\frac{1}{2}} \underline{w}_n + \lambda \left(\underline{w}_n^{*T} \underline{w}_n - 1 \right)$$
(6.26)

The optimization is therefore equivalent to finding the numerical radius of the transformed polarimetric interferometric matrix Π introduced in equation (6.20), as shown in equation (6.27):

$$\Pi = \sqrt{T^{-1}}\Omega_{12}\sqrt{T^{-1}} \Rightarrow \gamma_{opt} = r(\Pi)$$
(6.27)

There are many theorems and algorithms in the mathematics literature dealing with the concept of numerical radius (Murnaghan, 1932; Li, 1994; He, 1997; Mengi, 2005). Unfortunately there are no general analytical solutions available, but various numerical iterative algorithms have been proposed—one of which involves exactly the phase transformation and repeated eigensolution approach represented in equation (6.22). Furthermore, it leads to a third important approach to optimization, based not on coherence amplitude but on phase difference or coherence separation, as we now consider.

6.2.3 Maximum coherence separation and the coherence region

In the previous two sections we considered methods for finding the polarimetric scattering mechanisms \underline{w}_1 and \underline{w}_2 that maximize the interferometric coherence magnitude. Since the local phase variance in an interferogram is inversely proportional to coherence, this optimization will, by definition, lead to the interferogram with minimum phase noise. This important analytical result is somewhat marred by the practical issues of coherence bias, as discussed in Appendix 2 and demonstrated by example in Figure 6.2.

There is, however, a completely different approach to the optimization procedure. Instead of concerning ourselves with the local phase variance, we often seek a pair of scattering mechanisms w_1 and w_2 that maximize not the coherence amplitude but the *separation* of complex coherences in the complex plane. Physically these might then represent separated phase centres in a vegetation layer, for example (see Chapter 7) (Flynn, 2002; Tabb, 2002a). The first approach to this problem (Tabb, 2002a) was to develop an algorithm for maximising the phase separation, without regard to the coherence magnitude. However, this can cause problems when dealing with low coherence regions, as it can be sensitive to any noise in the data. A slightly modified approach is to consider the maximum separation of complex coherence values. This approach leads to a useful algorithm for application in the RVOG and related models (see Chapters 7 and 8). For this reason we consider this algorithm in more detail. This 'optimization of separation' can be conveniently formulated using the constrained approach where $\underline{w}_1 = \underline{w}_2$. In this case we found that the following eigenvalue equation could be used to maximize the real part of the complex coherence:

$$[T]^{-1}[\Omega_H(\phi)]\underline{w} = \lambda(\phi)\underline{w} \xrightarrow{\substack{\phi \\ \phi \\ \phi}} \Delta \tilde{\gamma}_{opt}$$
(6.28)

The desired maximum difference is then given by the maximum difference between eigenvalues of this matrix. Again we need to employ a phase transformation ϕ to ensure that we secure the global optimum separation. In this case we obtain a pair of \underline{w} vectors, \underline{w}_a and \underline{w}_b —one from each eigenvector corresponding to the max/min eigenvalues. The two complex coherences can then be explicitly evaluated, as shown in equation (6.29):

$$\left. \begin{array}{l} \tilde{\gamma}_{1} = \frac{\underline{\mathbf{w}}_{a}^{*T}[\Omega_{H}]\underline{w}_{a}}{\underline{\mathbf{w}}_{a}^{*T}[T]\underline{w}_{a}} \\ \tilde{\gamma}_{2} = \frac{\underline{\mathbf{w}}_{b}^{*T}[\Omega_{H}]\underline{w}_{b}}{\underline{\mathbf{w}}_{b}^{*T}[T]\underline{w}_{b}} \end{array} \right\} \Rightarrow \Delta \tilde{\gamma} = |\tilde{\gamma}_{1} - \tilde{\gamma}_{2}| = \Delta \tilde{\gamma}_{opt}$$
(6.29)

In summary, we have seen that there are three main approaches to coherence optimization in polarimetric interferometry:

- 1. The unconstrained amplitude optimization provides the most general mathematical solution, yielding the minimum phase variance interferogram across independent polarimetric variations at either end of the baseline.
- 2. The constrained amplitude approach yields a slightly sub-optimum solution, but one constrained to keep the polarimetry constant at either end of the baseline.
- 3. The constrained approach also yields a complex separation optimization to find the two scattering mechanisms with maximum interferometric separability inside the unit circle.

6.2.3.1 The coherence region

We can provide a useful geometrical interpretation of these various concepts using the coherence diagram. This is a unit circle representation of coherence in the complex plane (Figure 6.4). The first concept we can then consider is that of the coherence region inside this diagram (Flynn, 2002). For any given polarimetric interferometry matrix Λ_2 there will be some sub-region of the whole unit circle that encloses all possible values of coherence (for all states \underline{w}). This is called the coherence region of the matrix Λ_2 . We will see that in some cases this region may in fact shrink to a point, while in others it can include large parts of the circle. In general the shape and size of the region are determined by the nature of the scattering processes. We will see later, in Chapter 7, how to predict the limiting shape of the region for various canonical surface and volume scattering problems. Then, in Chapter 8, we will show how to use knowledge of the region shape to estimate important physical parameters (such as vegetation height) from radar data.

First we demonstrate how the boundary of the region can be computed numerically for the constrained case $(w_1 = w_2)$ using the eigenvalue equation derived



Fig. 6.4 Definition of the coherence region of a polarimetric interferometric matrix Λ_2

in equation (6.22). For each value of ϕ this eigenvalue equation yields the extreme values (through the maximum and minimum eigenvalues) of the real part of the coherence. For each of these eigenvalues there corresponds an eigenvector, which can be used to estimate a corresponding complex coherence, as shown in equation (6.30). These two coherences then define two points on the boundary of the coherence region, as shown schematically in Figure 6.4. Here we show an example elliptical coherence region (see equation (6.31)), and show, for a specified value of ϕ , how we obtain two samples of the boundary. By varying ϕ in the range $0 \le \phi \le \pi$ we can then reconstruct the whole boundary. This gives us a systematic way to visualize the boundary for an arbitrary coherency matrix Λ_2 .

$$[T]^{-1}[\Omega_{H}]\underline{w} = \lambda \underline{w} \begin{cases} [\Omega_{H}] = \frac{1}{2} \left(\Omega_{12} e^{i\phi} + \Omega_{12}^{*T} e^{-i\phi} \right) \\ [T] = \frac{1}{2} \left(T_{11} + T_{22} \right) \end{cases}$$

$$\overset{\lambda_{\max}, \underline{w}_{\max}}{\longrightarrow} \Rightarrow \begin{cases} \gamma_{\max}(\phi) = \frac{\underline{w}_{\max}^{*T} \Omega_{12} \underline{w}_{\max}}{\underline{w}_{\max}^{*T} T \underline{w}_{\max}} \\ \gamma_{\min}(\phi) = \frac{\underline{w}_{\min}^{*T} \Omega_{12} \underline{w}_{\min}}{\underline{w}_{\min}^{*T} T \underline{w}_{\min}} \end{cases}$$

$$(6.30)$$

As a more specific example, consider again the Λ_2 matrix shown in equation (6.15). This has a region dominated by three points: the three diagonal complex values of Ω_{12} . Figure 6.5 shows the corresponding triangular region for this matrix defined by the three diagonal values as vertices. As we vary the polarisation over all possible mechanisms the interferometric coherences will always be contained within this triangle. The boundaries of the region therefore define the various optimum values. In this case the constrained and unconstrained maximum amplitudes are equal to the white vertex. The separation optimization yields the white and black vertices, with a maximum phase difference of $\pi/2$.



Fig. 6.5 Example coherence region for example matrix in equation (6.15)

This provides a simple numerical example to illustrate the various optimization schemes. The real utility of these algorithms, however, is their application in the more general case, when the Ω_{12} matrix is full. Before considering such cases and their relationship to scattering theory, we first develop a useful subspace interpretation of the information contained in a full Ω_{12} matrix.

6.2.4 Subspace coherence region analysis: the SVD and Schur decompositions

The field of values concept (equation (6.23)) applies to arbitrary matrix dimension, but takes on a particularly simple form for 2×2 complex matrices. The field of values of any 2×2 matrix is an ellipse (for a formal proof see Murnaghan (1932)). More precisely, let a general 2×2 matrix *A* be defined as show in equation (6.31):

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \Rightarrow A = [U_2]^{*T} \begin{bmatrix} \lambda_1 & \delta \\ 0 & \lambda_2 \end{bmatrix} [U_2]$$
(6.31)

This matrix can, by Schur's theorem (see Appendix 1), always be written in terms of an upper diagonal form and a unitary matrix [U], as shown on the right-hand side of equation (6.31). Here λ_1 and λ_2 are the eigenvalues of **A**. It then follows that the field of values of **A** is an ellipse with two foci given by λ_1 and λ_2 and minor axis length $|\delta|$. The corresponding major axis length is given by $\sqrt{|\lambda_1 - \lambda_2|^2 + |\delta|^2}$. For the special case that $\delta = 0$ (in which case the matrix **A** is termed 'normal'—it can be diagonalized by a unitary transformation) we obtain a linear field of values, varying along a line stretching between the two eigenvalues. We shall see in Chapter 7 that such a limiting case plays an important role in the description of mixed surface and volume scattering.

There are clearly an infinite number of ways of generating such 2×2 matrices in polarimetric interferometry, simply by choosing a pair of polarisation vectors \underline{w}_X and \underline{w}_Y (and in the unconstrained version, a different pair \underline{w}_W and \underline{w}_Z for the other end of the baseline). These are then used to project the scattering vector data \underline{k} at each end of the baseline 1 and 2, which are used to generate a 4 × 4 projected (*p*) polarimetric interferometric coherency matrix, as shown (for the constrained case) in equation (6.32):

From this we can then generate a pre-whitened 2×2 matrix Π^p as shown in equation (6.33):

$$T = \frac{1}{2} \left(T_{11}^p + T_{22}^p \right) \Rightarrow \Pi^p = \sqrt{T^{-1}} \Omega_{12}^p \sqrt{T^{-1}}$$
(6.33)

The field of values of this matrix (for all projection vectors) is always an ellipse. The projection vectors can be chosen in different ways. One way is to use compact polarimetry (see Section 9.3.4), whereby a single transmit polarisation and (generally different) dual polarised receiver are used. A second is to employ physical modelling of the scene to isolate a subspace of polarisations where the desired phenomena (dihedral scattering in vegetation, for example) are isolated. However, a third important way is to start with the full Quadpol [*S*] matrix data, and then identify suitable subspaces by employing the Schur decomposition itself.

Our starting point for a general subspace analysis is the 3×3 pre-whitened polarimetric interferometric matrix Π as defined in equation (6.34). The singular vector (SVD) and Schur techniques may then be directly related to our two main approaches to optimization—the SVD suitable for an unconstrained approach to polarimetric interferometry, and the Schur for a constrained optimization approach. These ideas are summarized in equation (6.34):

$$\Pi = \sqrt{T^{-1}} \Omega_{12}^{p} \sqrt{T^{-1}} = \begin{cases} \begin{bmatrix} U_3 \end{bmatrix}^{*T} \begin{bmatrix} s_1 & 0 & 0 \\ 0 & s_2 & 0 \\ 0 & 0 & s_3 \end{bmatrix} \begin{bmatrix} V_3 \end{bmatrix} \text{ SVD} \\ \begin{bmatrix} U_3 \end{bmatrix}^{*T} \begin{bmatrix} \lambda_1 & \delta_{12} & \delta_{13} \\ 0 & \lambda_2 & \delta_{23} \\ 0 & 0 & \lambda_3 \end{bmatrix} \begin{bmatrix} U_3 \end{bmatrix} \text{ Schur}$$
(6.34)

In SVD we allow different vectors at either end of the baseline and obtain the singular values s_1 and so on, which we have shown in equation (6.20) correspond directly to optimum coherences. In the Schur approach, however, we constrain the decomposition to a single unitary matrix (corresponding to equal vectors at either end of the baseline), which leads to an upper triangular 3×3 matrix, also shown in equation (6.34). However, following the Schur approach we can now consider a set of 2×2 sub-matrices of this upper diagonal form in the knowledge that each will have an elliptical coherence region as discussed above. For example, we can consider the subspace formed by the pairings 1,2, 1,3 or 2,3 in equation (6.34). This can be useful if, for example, there is noise in part of the subspace we wish to remove, or if we are seeking the subspace with the most linear coherence region based on physical modelling such as RVOG (see Section 7.4.2).

We can make the link between this approach and the general projection ideas of equation (6.32) by noting that the unitary matrix [U] obtained in the Schur decomposition can be written as a set of three column vectors, \underline{u} , corresponding to projection vectors \underline{w} by a basis transformation, as shown in equation (6.35):

$$[U_3] = \begin{bmatrix} \underline{u}_1 & \underline{u}_2 & \underline{u}_3 \end{bmatrix} \Rightarrow \underbrace{w_1}_2 = T^{-\frac{1}{2}} \underbrace{u_1}_2 \qquad (6.35)$$
$$\underbrace{w_3}_3 = T^{-\frac{1}{2}} \underbrace{u_3}_3$$

By setting the pair x,y equal to 1,2, 1,3, and 2,3 in equation (6.32), we then provide a link between the general Schur decomposition and projection approach. Note that for each pairing we can directly calculate the shape of the coherence region analytically, since it is always elliptical, with two foci λ_x and λ_y , minor axis length $|\delta_{xy}|$ and major axis length $\sqrt{|\lambda_x - \lambda_y|^2 + |\delta_{xy}|^2}$. As a simple example, consider again the matrix shown in equation (6.15), with a region illustrated in Figure 6.5. In this case the Π matrix has the following simple form:

$$\Pi = \begin{bmatrix} 0.9 & 0 & 0\\ 0 & 0.6e^{i\frac{\pi}{4}} & 0\\ 0 & 0 & 0.4e^{i\frac{\pi}{2}} \end{bmatrix}$$
(6.36)

Here the three subspace regions reduce to line segments joining the pairs of eigenvalues, as shown in Figure 6.5. We now return to the issue of numerical bias in the context of these new coherence optimization techniques.

6.2.5 Numerical bias in coherence optimization

In the previous section we developed some useful analytical results concerning the issue of coherence optimization and its impact on determining the dynamic range of interferometric coherence variation with polarisation. In this section we briefly consider a practical issue, to be considered when applying these ideas to measured radar data: the impact of coherence bias in matrix estimation, and how it impacts on estimation of the optimum coherences (Touzi, 1999).

In practice we often have no knowledge of the detailed form of the coherency matrix Λ_2 , and must instead estimate it from experimental data. Adopting a maximum likelihood approach, estimates can be made for the three sub-matrices involved by averaging the scattering vector data, as shown in equation (6.37):

$$\hat{T}_{11} = \frac{1}{L} \sum_{i=1}^{L} \underline{k}_{1i} \underline{k}_{1i}^{*T}, \quad \hat{T}_{22} = \frac{1}{L} \sum_{i=1}^{L} \underline{k}_{2i} \underline{k}_{2i}^{*T}, \quad \hat{\Omega}_{12} = \frac{1}{L} \sum_{i=1}^{L} \underline{k}_{1i} \underline{k}_{2i}^{*T} \quad (6.37)$$

The question now is, what is the influence of the number of samples ('looks' in radar imaging terms) L on the coherence optimization algorithms? As $L \rightarrow \infty$ we should obtain the true matrices (since in this limit the matrices converge to their correct values), but for small L we can expect overestimation of coherence and hence distortion of the coherence region. To analytically study the effects of L on optimum coherence estimation is a difficult task (see Touzi, 1999; Lopez-Martinez, 2005), and here we therefore employ some illustrative numerical simulations based on use of Monte Carlo simulations (see Appendix 3 for details) to illustrate the nature of the problems involved, and to form some general conclusions about bias effects in optimization methods. We again make use of the random volume scattering example shown in equation (6.15), and this time use the simulated data to estimate the matrices using equation (6.37) *before* applying the various constrained and unconstrained optimization algorithms.

Figure 6.6 shows the results of applying the unconstrained optimization algorithm to the estimated matrices (again each point formed from an average of 256 coherence estimates). We note that the bias issues are more severe than for the standard coherence estimation (shown dashed, for reference, in Figure 6.6). This reflects the underlying higher dimensionality of the general unconstrained optimization process. Direct coherence estimation implies that we have *a priori* knowledge of the w vectors—in this case just the Pauli scattering vectors—and so can project *before* we undertake the coherence estimate to obtain the improved convergence shown in the dashed lines in Figure 6.6. However, for the unconstrained optimization process we not only have to estimate coherence values but also do not know the projection vectors themselves. These too must be estimated from the data. Hence we need to estimate a larger number of parameters from the data itself. This increased dimensionality requires an increased number of looks for convergence. This provides a qualitative explanation of the increased bias seen in Figure 6.6.



Fig. 6.6 Coherence bias in matrix estimation of optimum coherence triplet in equation (6.15)



Fig. 6.7 Distortion of coherence region of equation (6.15) arising from coherence bias in matrix estimation method

We note also that the triplet of optimum states (the three eigenvalues of [K]) have different bias issues. The first and second optima are overestimated for small L, but the smallest optimum value is actually underestimated. This correlated bias behaviour between eigenvalues means that the apparent dynamic range of coherence variability with polarisation is overestimated for small L. We see that it takes in excess of L = 50 looks for the estimate bias to settle down, but note again slow convergence even beyond this point.

This overestimation of dynamic range is also apparent in the estimation of the coherence region. If we employ coherency matrix estimates for L = 6 looks and then L = 50 looks we obtain typical region estimates as shown in Figure 6.7. Here, in grey we show the estimated boundary region for L = 6, and note its overestimation compared to the true region (shown as the black triangle). For L = 50 we see a much better estimate (shown in black), more accurately reflecting the bounds of the true coherence region, and hence prociding a better estimate of the true dynamic range.

We now turn to consider how the physical structure of surface and volume scattering controls the size and shape of the coherence region.

7

The coherence of surface and volume scattering

In this chapter we investigate in more detail the shape and structure of the limiting form of the coherence region for vector surface and volume scattering problems. In Section 6.2.3.1 the coherence region is defined as the region in the complex plane bounding the variation of interferometric coherence with polarisation. Here we extend this idea to define a related concept: the coherence loci, defined as the curves traced out by variation of interferometric coherence with physical parameters of a scattering model (Papathanassiou, 2001; Cloude, 2003). Our objective is to relate the coherence loci as a limiting form of the coherence region (as the number of looks tends to infinity) in order to establish strategies in Chapter 8 for using polarimetric interferometry for physical parameter estimation.

We begin by looking at simple models of surface and volume-only scattering. We then consider extension of these ideas to multilayer media which, importantly, will allow us to consider *combinations* of surface and volume effects. We then look in detail at two important models widely used in the literature for interpreting coherence diagram: the random-volume-over-ground, or RVOG (Treuhaft, 1996, 2000a; Papathanassiou, 2001; Cloude, 2003), which is closely related to an interferometric version of the water cloud model IWCM (Askne, 1997, 2003, 2007) (see Section 3.5.1); and the oriented-volume-over-ground, or OVOG (Treuhaft, 1999; Cloude, 2000a). Both these models are characterized by having a small number of independent physical parameters, often fewer than observables in the scattered field, so enabling consideration of methods for estimation of these parameters from data (see Chapter 8).

As we shall see, both of these models (RVOG and OVOG) make assumptions about the vertical variation of scattering in the layered media (through the structure function), which naturally leads us to consider a more general approach termed coherence tomography (Cloude, 2006b, 2007a) that permits arbitrary structure function and allows an efficient parameterization of the dependence of coherence of changes in structure.

In general terms we note that the coherence loci must somehow be related to variation of the vertical structure function f(z) with polarisation. For example, if the scatterers in a scene do not change *relative* amplitude as \underline{w} changes, then the structure function, whatever shape it has, will be constant, and the coherence will be constant with polarisation, so yielding a point coherence loci. This point can then be stretched to a radial line by adding polarisation-dependent temporal, or SNR, decorrelation, but the underlying physics will be determined by a point in the complex diagram. We now investigate this relationship between coherence loci shape and structure function variations in more detail, for surface and volume scattering scenarios.

7.1 Coherence loci for surface scattering

The first issue we face is in defining surface scattering in the context of interferometry. By definition, surface scattering occurs at a discontinuity between two media, and hence a good model for its vertical structure function would be a Dirac delta function located at the interface between the media, as shown schematically in Figure 7.1, where the surface is clearly located at position $z = z_0$. However, we have seen that in microwave remote sensing of natural surfaces there is always some penetration of the wave into the lower medium, depending on the effective dielectric constant (see Section 3.1.1.1), and hence it is of interest to consider the circumstances under which this delta function assumption is supported. In the context of interferometry, what is important is not so much the absolute penetration depth but its value scaled to β_z —the vertical sensitivity of the interferometer. Hence by combining the definition of penetration depth δ_p (equation (3.12)) with the baseline dependence of vertical wavenumber β_z , we can obtain a relationship between effective complex material constant $\varepsilon_r = \varepsilon' - i\varepsilon''$ and baseline geometry, as shown in equation (7.1). Here we set a threshold of 0.1 radians for the product, as this represents a typical interferometric phase shift due to wave penetration of only 5° and a maximum volume decorrelation of only 0.998. These are within the bounds of estimation error for typical interferometer geometries, and thus represent a somewhat arbitrary but realistic threshold for the delta function assumption to apply.

$$\frac{\beta_{z}\delta_{p}}{2} < 0.1$$

$$\Rightarrow \frac{4\pi\Delta\theta}{\lambda\sin\theta} \cdot \frac{\lambda}{2\pi} \frac{\sqrt{\varepsilon'}}{\varepsilon''} = \frac{2\Delta\theta}{\sin\theta} \frac{\sqrt{\varepsilon'}}{\varepsilon''} \approx \frac{2B_{\perp}}{R_{o}\sin\theta} \frac{\sqrt{\varepsilon'}}{\varepsilon''} < 0.1$$

$$\Rightarrow \frac{\sqrt{\varepsilon'}}{\varepsilon''} < \frac{R_{o}\sin\theta}{20B_{\perp}}$$
(7.1)

If this inequality is satisfied (and we further assume that range spectral filtering has been employed to remove any baseline decorrelation (see Section 5.1.1.1)), the interferometric coherence can then be estimated as shown in equation (7.2), where we have also, for the moment, ignored any temporal or SNR decorrelation terms.

$$\hat{\gamma} = e^{i\beta_{z}z_{o}} \frac{\int_{o}^{h_{v}} \delta\left(z\right) e^{i\beta_{z}z} dz}{\int_{o}^{h_{v}} \delta\left(z\right) dz} = e^{i\beta_{z}z_{o}} = e^{i\phi_{o}}$$
(7.2)

We see a simple result, with a coherence of unity and phase depending on the surface elevation. Turning now to the polarisation dependence, as we vary polarisation so the backscatter amplitude from the surface will change. We can propose this variation of backscatter as a reflection symmetric depolariser (see Section 2.4.2.3), which has a polarimetric coherency matrix as shown in equation (7.3). In the absence of temporal and SNR decorrelation we can now calculate the optimum coherence values from the corresponding [*K*] matrix, as



Fig. 7.1 Idealized vertical structure function for surface scattering

shown in equation (7.3):

$$K = T_{11}^{-1} \Omega_{12} T_{11}^{-1} \Omega_{12}^{*T}$$

$$= \begin{bmatrix} t_{11} & t_{12} & 0 \\ t_{12}^* & t_{22} & 0 \\ 0 & 0 & t_{33} \end{bmatrix}^{-1} e^{i\phi_0} \begin{bmatrix} t_{11} & t_{12} & 0 \\ t_{12}^* & t_{22} & 0 \\ 0 & 0 & t_{33} \end{bmatrix}^{-1} e^{-i\phi_0} \begin{bmatrix} t_{11} & t_{12} & 0 \\ t_{12}^* & t_{22} & 0 \\ 0 & 0 & t_{33} \end{bmatrix}^{-1} e^{-i\phi_0} \begin{bmatrix} t_{11} & t_{12} & 0 \\ t_{12}^* & t_{22} & 0 \\ 0 & 0 & t_{33} \end{bmatrix}^{-1} e^{-i\phi_0} \begin{bmatrix} t_{11} & t_{12} & 0 \\ t_{12}^* & t_{22} & 0 \\ 0 & 0 & t_{33} \end{bmatrix}^{-1} e^{-i\phi_0} \begin{bmatrix} t_{11} & t_{12} & 0 \\ t_{12}^* & t_{22} & 0 \\ 0 & 0 & t_{33} \end{bmatrix}^{-1} e^{-i\phi_0} \begin{bmatrix} t_{11} & t_{12} & 0 \\ t_{12}^* & t_{22} & 0 \\ 0 & 0 & t_{33} \end{bmatrix}^{-1} e^{-i\phi_0} \begin{bmatrix} t_{11} & t_{12} & 0 \\ t_{12}^* & t_{22} & 0 \\ 0 & 0 & t_{33} \end{bmatrix}^{-1} e^{-i\phi_0} \begin{bmatrix} t_{11} & t_{12} & 0 \\ t_{12}^* & t_{22} & 0 \\ 0 & 0 & t_{33} \end{bmatrix}^{-1} e^{-i\phi_0} \begin{bmatrix} t_{11} & t_{12} & 0 \\ t_{12}^* & t_{22} & 0 \\ 0 & 0 & t_{33} \end{bmatrix}^{-1} e^{-i\phi_0} \begin{bmatrix} t_{11} & t_{12} & 0 \\ t_{12}^* & t_{22} & 0 \\ 0 & 0 & t_{33} \end{bmatrix}^{-1} e^{-i\phi_0} \begin{bmatrix} t_{11} & t_{12} & 0 \\ t_{12}^* & t_{22} & 0 \\ 0 & 0 & t_{33} \end{bmatrix}^{-1} e^{-i\phi_0} \begin{bmatrix} t_{11} & t_{12} & 0 \\ t_{12}^* & t_{22} & 0 \\ 0 & 0 & t_{33} \end{bmatrix}^{-1} e^{-i\phi_0} \begin{bmatrix} t_{11} & t_{12} & 0 \\ t_{12}^* & t_{22} & 0 \\ 0 & 0 & t_{33} \end{bmatrix}^{-1} e^{-i\phi_0} \begin{bmatrix} t_{11} & t_{12} & 0 \\ t_{12}^* & t_{22} & 0 \\ 0 & 0 & t_{33} \end{bmatrix}^{-1} e^{-i\phi_0} \begin{bmatrix} t_{11} & t_{12} & 0 \\ t_{12}^* & t_{22} & 0 \\ 0 & 0 & t_{33} \end{bmatrix}^{-1} e^{-i\phi_0} \begin{bmatrix} t_{11} & t_{12} & 0 \\ t_{12}^* & t_{22} & 0 \\ 0 & 0 & t_{33} \end{bmatrix}^{-1} e^{-i\phi_0} \begin{bmatrix} t_{11} & t_{12} & 0 \\ t_{12}^* & t_{22} & 0 \\ 0 & 0 & t_{33} \end{bmatrix}^{-1} e^{-i\phi_0} \begin{bmatrix} t_{11} & t_{12} & 0 \\ t_{12}^* & t_{22} & 0 \\ 0 & 0 & t_{33} \end{bmatrix}^{-1} e^{-i\phi_0} \begin{bmatrix} t_{11} & t_{12} & 0 \\ t_{12}^* & t_{23} & t_{24} \end{bmatrix}^{-1} e^{-i\phi_0} \begin{bmatrix} t_{11} & t_{12} & t_{13} & t_{13} \end{bmatrix}^{-1} e^{-i\phi_0} \begin{bmatrix} t_{11} & t_{12} & t_{13} & t_{13} \\ t_{13}^* & t_{13} & t_{13} \end{bmatrix}^{-1} e^{-i\phi_0} \begin{bmatrix} t_{11} & t_{12} & t_{13} & t_{13} \\ t_{13}^* & t_{13} & t_{13} & t_{13} & t_{13} \end{bmatrix}^{-1} e^{-i\phi_0} \begin{bmatrix} t_{11} & t_{13} & t_{13} & t_{13} \\ t_{13}^* & t_{13} & t_{13} & t_{13} & t_{13} \end{bmatrix}^{-1} e^{-i\phi_0} \begin{bmatrix} t_{13} & t_{13} & t_{13} & t_{13} & t_{13} & t_{13} \\ t_{13}^* & t_{1$$

Here we see that [K] is just the identity matrix, indicating an interferometric coherence of unity for *all* polarisation states. This is just a consequence of the fact that although the absolute level of the delta function in the structure function f(z) varies with changes in polarisation, its position remains fixed at the surface boundary $z = z_0$, and hence the coherence remains the same for all polarisations. In this case the corresponding coherence region shrinks to a point on the circumference of the unit circle (the angular position of which depends on $\beta_z z_0$).

In practice this result must be extended to include variations in SNR with polarisation. The backscatter power from smooth surfaces can be very low, especially in the crosspolarised channel, and this variation will be apparent as a polarisation-dependent coherence, as shown in equation (7.4):

$$s/n = \frac{\underline{w}^{*T} [T_{11}] \underline{w}}{n} \Rightarrow \gamma_{snr}(\underline{w}) = \frac{1}{1 + \frac{n}{w^{*T} [T_{11}] \underline{w}}}$$
(7.4)

where the noise power *n* can be estimated directly in reciprocal backscatter problems as the smallest eigenvalue of the HV/VH N = 2 coherency matrix, as shown in equation (7.5) (Hajnsek, 2001):

$$n = \frac{1}{2} \left(\left\langle S_{HV} S_{HV}^* \right\rangle + \left\langle S_{VH} S_{VH}^* \right\rangle - \sqrt{\left(\left\langle S_{HV} S_{HV}^* \right\rangle - \left\langle S_{VH} S_{VH}^* \right\rangle \right)^2 + 4 \left\langle S_{HV} S_{VH}^* \right\rangle \left\langle S_{VH} S_{HV}^* \right\rangle} \right)$$
(7.5)

Temporal decorrelation can, of course, occur with surface changes between passes in repeat-pass interferometry. However, there is no strong reason why such changes should occur in a polarisation-sensitive way, and so we can realistically model such effects as a scalar multiplier applied equally to all polarisation channels. In this way our final expression for the polarisation dependence of coherence in surface scattering scenarios takes the form shown in equation (7.6):

$$\hat{\gamma} = \gamma_t \gamma_{SNR}(\underline{w}) e^{i\phi_o} \tag{7.6}$$

This corresponds to a coherence loci given by a radial line segment in the complex plane, as shown schematically in Figure 7.2. Note that although the coherence amplitude can vary with polarisation, the average phase of the coherence is constant, geometrically implied by the radial nature of the coherence loci in the coherence plane. In this case the maximum coherence corresponds to the polarisation with maximum signal-to-noise ratio, representing one end of the loci, as shown in Figure 7.2. The other boundary of the loci corresponds



Fig. 7.2 Coherence loci for surface scattering: ideal case (left), and with SNR and temporal decorrelation (right)

similarly to the polarisation with minimum SNR. The whole line position (and scaled line length) is dictated by the temporal decorrelation γ_t . At extremes, when $\gamma_t = 1$, the line maximum can approach the unit circle at the point ϕ_0 as shown. When $\gamma_t = 0$, the whole line reduces to a point at the origin.

7.2 Coherence loci for random volume scattering

We now turn to consider determination of the coherence region for volume scattering. We begin with the strongest polarisation symmetry assumption: azimuthal symmetry, which leads to a random-volume approximation, for which the polarisation coherency matrix is diagonal and of the general form shown in equation (7.7):

$$\langle [T] \rangle = m_{\rm v} \begin{bmatrix} 1 & 0 & 0 \\ 0 & s & 0 \\ 0 & 0 & s \end{bmatrix} \quad 0 \le s \le 1.0 \tag{7.7}$$

where the absolute scattering cross-section m_v is given, for example, by the water cloud model (see Section 3.5.1), and *s* depends on particle shape and varies for single scattering in the range 0.5 (dipole cloud) to 0 (spheres). This strong symmetry assumption also has an important impact on the shape of the coherence loci, as we now demonstrate.

Ignoring, for the moment, temporal and SNR effects, we can calculate the optimum coherences from the [K] matrix, as shown in equation (7.8):

$$K = T_{11}^{-1} \Omega_{12} T_{11}^{-1} \Omega_{12}^{*T}$$

$$= \frac{1}{I_1} \begin{bmatrix} \frac{1}{t_{11}} & 0 & 0\\ 0 & \frac{1}{t_{22}} & 0\\ 0 & 0 & \frac{1}{t_{33}} \end{bmatrix} I_2 \begin{bmatrix} t_{11} & 0 & 0\\ 0 & t_{22} & 0\\ 0 & 0 & t_{33} \end{bmatrix} \frac{1}{I_1} \begin{bmatrix} \frac{1}{t_{11}} & 0 & 0\\ 0 & \frac{1}{t_{22}} & 0\\ 0 & 0 & \frac{1}{t_{33}} \end{bmatrix} I_2^* \begin{bmatrix} t_{11} & 0 & 0\\ 0 & t_{22} & 0\\ 0 & 0 & \frac{1}{t_{33}} \end{bmatrix} I_2^* \begin{bmatrix} t_{11} & 0 & 0\\ 0 & t_{22} & 0\\ 0 & 0 & \frac{1}{t_{33}} \end{bmatrix} I_2^* \begin{bmatrix} t_{11} & 0 & 0\\ 0 & t_{22} & 0\\ 0 & 0 & t_{33} \end{bmatrix} I_1^* \begin{bmatrix} t_{11} & 0 & 0\\ 0 & t_{22} & 0\\ 0 & 0 & t_{33} \end{bmatrix} I_1^* I_2^* I_2^* I_1^* I_2^* I_2^* I_1^* I_2^* I_1^* I_2^* I_1^* I_2^* I_1^* I_2^* I_2^* I_1^* I_2^* I_1^* I_1^* I_2^* I_1^* I_1^* I_2^* I_1^* I_2^* I_1^* I_2^* I_1^* I_2^* I_1^* I_1^* I_2^* I_1^* I_1^* I_2^* I_1^* I_2^* I_1^* I_2^* I_1^* I_1^$$



Fig. 7.3 Schematic representation of an arbitrary vertical structure function

We note that [K] is again a multiple of the identity matrix, indicating that the coherence does not change with polarisation and, as we found for surface scattering, the coherence loci reduces to a point in the coherence diagram. However, unlike surface scattering, the point does not lie on the unit circle. Instead it lies within the circle at a point determined by the complex volume decorrelation caused by the structure function f(z) (see Section 5.2.4). The integral factors I_1 and I_2 in equation (7.8) can be expressed in terms of the Legendre expansion of the structure function, as shown in equation (7.9). The key consequence of the random symmetry assumption is that the Legendre coefficients are independent of polarisation, and so the structure function, which can be arbitrary, as shown in Figure 7.3, must remain invariant to changes in polarisation.

$$I_{2} = e^{i\beta_{z}z_{o}} \int_{0}^{h_{v}} f(z')e^{i\beta_{z}z'}dz' \\ I_{1} = \int_{0}^{h_{v}} f(z')dz' \\ = e^{i\beta_{z}z_{0}}e^{i\frac{\beta_{z}h_{v}}{2}}\frac{(1+a_{0})f_{0}+a_{1}f_{1}+a_{2}f_{2}+\dots+a_{n}f_{n}}{(1+a_{0})}$$
(7.9)

The effect of signal-to-noise ratio will be similar to that found for surfaces; that is, to provide a polarisation-sensitive radial shift of the coherence towards the origin. However, since volume scattering is generally more depolarising than surface scattering, the variation of scattered power with polarisation will be less, and hence SNR effects less important, than they are for surfaces. On the other hand, temporal decorrelation can be much more important for volume scattering, especially in vegetation applications, due to its susceptibility to wind-driven motion on short time-scales. To further complicate issues, the effects of temporal changes may not be uniform across the structure function. For example, wind-blown motion may affect the top of the vegetation layer more than the lower regions. To accommodate this we can modify the coherence integrals to include in the numerator (I_2) a new temporal structure function g(z), as shown in equation (7.10):

$$I_{2} = e^{i\beta_{z}z_{o}} \int_{0}^{h_{v}} g(z')f(z')e^{i\beta_{z}z'}dz' \\I_{1} = \int_{0}^{h_{v}} f(z')dz'$$
 $\Rightarrow \tilde{\gamma} = \frac{I_{2}}{I_{1}}$ (7.10)

The function g(z) will vary between 0 and 1, being zero in regions of maximum change and 1 for zero change. In terms of the Legendre expansion, g(z) will of course have its own expansion coefficients, which in general will be different from those of f(z), and hence the effect of temporal decorrelation must formally be evaluated as a product of Legendre series in the numerator I_2 . In the simplest case (and the one most often used in the literature) we can assume that $g(z) = \gamma_t$ —a constant function with height, in which case the overall coherence can



Fig. 7.4 Coherence loci for random volume scattering: ideal case (left), and with combined SNR and temporal decorrelation (right)

be expressed as shown in equation (7.11):

$$\tilde{\gamma} = \gamma_{SNR}(\underline{w})\gamma_t e^{i\beta_z z_0} e^{i\frac{\beta_z h_v}{2}} \frac{(1+a_0)f_0 + a_1f_1 + a_2f_2 + \dots + a_nf_n}{(1+a_0)}$$
(7.11)

Figure 7.4 summarizes the coherence loci for random volume scattering. Again, as for surface scattering, it is represented by a radial line in the complex coherence diagram.

There are two important differences between the surface and volume coherence regions. The maximum coherence in the surface case (in the absence of temporal and SNR effects) was on the unit circle (the point on the left of Figure 7.3). However, in volume scattering, even in the limit that $\gamma_t = \gamma_{snr} = 1$, the maximum coherence no longer lies on the unit circle but somewhere inside the exact location depending on the baseline geometry and importantly on the structure function of the volume scattering. The second difference is the presence of phase bias in the volume scattering case. The phase of the coherence does not correspond to the bottom of the layer, but is offset by a term ϕ_b , the value of which depends on the structure function. These observations provide us with our first important link between coherence and important structural parameters. We now consider two special cases: first the exponential structure function, and then issues related to orientation effects in the volume.

7.2.1 Special case I: the exponential profile

We have seen that under azimuthal polarisation symmetry, the structure function for random volume scattering can be arbitrary, as long as it the same in all polarisations. However, one special case is of interest because of its relation to physical models of propagation through an homogeneous layer. This is the exponential profile, used in deriving the water cloud model (WCM) for backscatter in Section 3.5.1 and in the study of volume decorrelation in Section 5.2.4.2. In this case we can evaluate the complex coherence explicitly, without the need for a Legendre expansion, as shown in equation (7.12). We note that the volume decorrelation is now a function of just two physical parameters: the height of the layer (h_v), and the mean extinction σ_e . This gives us two physical parameters to locate the coherence point in the complex plane. As this point is specified by two measurements (amplitude and phase) there is a good match between observables and unknown parameters. However, the match is spoilt by the addition of the unknown phase of the surface $\phi(z_0)$. This acts essentially as a new physical parameter (the location of the bottom of the layer) that must also be estimated from the data. Hence we now have three unknowns and only two observations. Nonetheless, this concept of reducing the number of parameters required to describe the structure function so as to better match the number of observations is an important one. We shall see in the case of layered media that it leads us to a convenient solution for estimation of physical parameters from data.

$$\begin{split} \tilde{\nu}(\underline{w}) &= \gamma_{SNR}\left(\underline{w}\right) \gamma_{t} \frac{I_{2}}{I_{1}} \\ &= \gamma_{SNR}\left(\underline{w}\right) \gamma_{t} \frac{e^{-\frac{2\alpha_{e}h_{v}}{\cos\theta_{o}}} \int_{0}^{h_{v}} e^{\frac{2\alpha_{e}z'}{\cos\theta_{o}}} e^{i\beta_{z}z'} dz'}{e^{-\frac{2\alpha_{e}h_{v}}{\cos\theta_{o}}} \int_{0}^{h_{v}} e^{\frac{2\alpha_{e}z'}{\cos\theta_{o}}} dz'} \\ &= \gamma_{SNR}\left(\underline{w}\right) \gamma_{t} \frac{2\sigma_{e} e^{i\phi(z_{o})}}{\cos\theta_{o}(e^{2\sigma_{e}-h_{v}/\cos\theta_{o}}-1)} \int_{0}^{h_{v}} e^{i\beta_{z}z'} e^{\frac{2\alpha_{e}-z'}{\cos\theta_{o}}} dz' \\ &= \gamma_{SNR}\left(\underline{w}\right) \gamma_{t} \frac{p}{p_{1}} \frac{e^{p_{1}h_{v}}-1}{e^{ph_{v}}-1} \text{ where } \begin{cases} p = \frac{2\sigma_{e}}{\cos\theta} \\ p_{1} = p + i\beta_{z} \\ \beta_{z} = \frac{4\pi\Delta\theta}{\lambda\sin\theta} \approx \frac{4\pi B_{n}}{\lambda H\tan\theta} \end{cases} \\ &= \gamma_{SNR}\left(\underline{w}\right) \gamma_{t} \tilde{\gamma}_{v} \end{split}$$
(7.12)

One important form of the exponential structure approximation occurs when we let the depth of the layer tend to infinity. In this case, taking the limits of equation (7.12), and ignoring for the moment temporal and SNR decorrelation, we obtain the following special form of the coherence for an infinitely thick half-space.

$$\tilde{\gamma} = e^{i\beta_z z_o} \frac{p}{p_1} \frac{e^{p_1 h_v} - 1}{e^{p h_v} - 1} \begin{cases} p = \frac{2\sigma_e}{\cos\theta} \\ p_1 = p + i\beta_z \\ \beta_z = \frac{4\pi\Delta\theta}{\lambda\sin\theta} \approx \frac{4\pi B_n}{\lambda H \tan\theta} \end{cases} \stackrel{\lim h_v \to \infty}{\longrightarrow} \tilde{\gamma} = e^{i\beta_z z_o} \frac{p}{p_1} e^{i\beta_z h_v} \end{cases}$$
$$\Rightarrow \tilde{\gamma} e^{-i\beta_z (z_o + h_v)} = \frac{1}{1 + i\frac{\beta_z \cos\theta}{2\sigma_e}} \tag{7.13}$$

In this case it makes more sense to shift the phase origin to the top of the volume rather than the lower, in which case we obtain the following expression for the complex coherence:

$$\tilde{\gamma}(h_{\infty}) = \frac{1}{1 + i\frac{\beta_c \cos\theta}{2\sigma_r}}$$
(7.14)

This is a function of only one physical parameter: the mean extinction. The coherence loci for this model has a simple form, forming a semicircle in the coherence plane, starting on the unit circle at the top phase reference for infinite extinction (set to zero phase for convenience in Figure 7.5), and moving



Fig. 7.5 Coherence loci for semi-infinite random volume scattering medium with varying extinction

towards the origin (zero coherence) for zero extinction, as shown by the line in Figure 7.5.

This represents one of the simplest possible models for volume decorrelation, and when combined with a measured coherence (shown as the point in Figure 7.5), provides a means for estimation of the mean extinction in the volume from the coherence magnitude of a single polarisation interferogram. (We saw in the water cloud model (WMC), in Section 3.5.1, that this extinction is often directly related to the water content $m_{\rm y}$, and hence we can often use this extinction as a proxy for water content.) However, the assumption of an infinite depth restricts this approach to applications where layer thickness greatly exceeds wave penetration depth. Important examples are thick land-ice (Dall, 2003; Sharma, 2007) and high-frequency penetration of vegetation and snow; but in general terms the assumptions of this model are not robust, and layer thicknesses are often small compared to penetration, so that scattering from the underlying bounding surface cannot be ignored. Treatment of these scenarios will require multilayer scattering models to be developed in the next section, but first we consider an important variation on the exponential structure function assumption: the case of oriented volume scattering.

7.2.2 Special case II: oriented volume scattering

In many agricultural crop and ice remote sensing problems, the scatterers in a volume may have residual orientation correlation due to their natural structure (stalks in a wheat field, for example). The propagation of signals through such a volume can no longer be assumed to be isotropic. Clearly, polarisations parallel and perpendicular to the mean orientation axis will suffer different extinctions. We considered a coherency matrix formulation of such propagation effects in Section 4.2.6, and noted that essentially we again need to make an exponential structure function approximation through such volumes, but now one where the

exponential coefficient itself varies as a function of polarisation. In this section we consider the coherence loci for such oriented volume scattering (Treuhaft, 1999; Cloude, 2000a; Ballester-Berman, 2005, 2007).

In such cases the volume has two eigenpolarisation propagation states \underline{x} and \underline{y} (which for an homogeneous channel (see Section 1.2.7) will be orthogonal). Only along these eigenpolarisations is the propagation simple, in the sense that the polarisation state does not change with penetration into the volume. If, however, there is some mismatch between the wave polarisation coordinates and the medium's eigenstates, then there arises a complicated situation in which the polarisation of the incident field changes as a function of distance into the volume. Here we demonstrate that the coherence optimizer always obtains a matched solution, and is thus useful in the application of parameter estimation schemes to oriented volume scattering problems. It also leads to determination of the coherence loci for such cases.

Essentially we now assume that the medium has backscatter reflection symmetry about the (unknown) axis of its eigenpolarisations (rather than azimuthal symmetry as in the random volume case), and so we obtain a polarimetric coherency matrix [T], and from this the covariance matrix [C], for backscatter (using the unitary transformation as derived in equation (7.16)), as shown in equation (7.15):

$$[T] = \begin{bmatrix} t_{11} & t_{12} & 0\\ t_{12}^* & t_{22} & 0\\ 0 & 0 & t_{33} \end{bmatrix} \xrightarrow{[C] = [U_{LP3}][T][U_{LP3}]^{-1}} [C] = \begin{bmatrix} c_{11} & 0 & c_{13}\\ 0 & c_{22} & 0\\ c_{13}^* & 0 & c_{33} \end{bmatrix}$$
(7.15)

$$[U_{LP3}] = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 & 0\\ 0 & 0 & \sqrt{2}\\ 1 & -1 & 0 \end{bmatrix}$$
(7.16)

We can then also relate the [K] optimization matrices from equation (6.14) in the two representations, as shown in equation (7.17):

$$K_{T} = T_{11}^{-1} \Omega_{12} T_{11}^{-1} \Omega_{12}^{*T}$$

$$\Rightarrow K_{C} = (U_{LP3} T_{11}^{-1} U_{LP3}^{-1}) (U_{LP3} \Omega_{12} U_{LP3}^{-1}) (U_{LP3} T_{11}^{-1} U_{LP3}^{-1}) (U_{LP3} \Omega_{12}^{*T} U_{LP3}^{-1})$$

$$\Rightarrow K_{C} = U_{LP3} K_{T} U_{LP3}^{-1} \Leftrightarrow K_{T} = U_{LP3}^{-1} K_{C} U_{LP3}$$
(7.17)

We make one further assumption: that the eigenpolarisations \underline{x} and \underline{y} are orthogonal linear states; but we do allow for a mismatch in the angle between these states and the radar coordinates by an angle ψ . We can now obtain an expression for the coherency matrix $[T_{11}] = [T_{22}]$ and interferometry matrix $[\Omega_{12}]$ for an oriented volume extending from $z = z_0$ to $z = z_0 + h_v$ as vector volume integrals shown in equations (7.18) and (7.19) (see equation (4.69):

$$[\Omega_{12}] = e^{i\phi(z_o)}R(2\psi) \left\{ \int_0^{h_v} e^{i\beta_z z'} e^{\frac{(\sigma_x + \sigma_y)z'}{\cos\theta_o}} P(\tau) TP(\tau^*) dz' \right\} R(-2\psi) \quad (7.18)$$

$$[T_{11}] = R\left(2\psi\right) \left\{ \int_0^{h_v} e^{\frac{(\sigma_x + \sigma_y)z'}{\cos\theta_o}} P\left(\tau\right) TP(\tau^*) dz' \right\} R\left(-2\psi\right)$$
(7.19)

where for clarity we have dropped the brackets around matrices and define the following terms:

$$R(\psi) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos \psi & \sin \psi \\ 0 & -\sin \psi & \cos \psi \end{bmatrix}$$
(7.20)
$$P(\tau)TP(\tau^*) = \begin{bmatrix} \cosh \tau & \sinh \tau & 0 \\ \sinh \tau & \cosh \tau & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} t_{11} & t_{12} & 0 \\ t_{12}^* & t_{22} & 0 \\ 0 & 0 & t_{33} \end{bmatrix}$$
$$\times \begin{bmatrix} \cosh \tau^* & \sinh \tau^* & 0 \\ \sinh \tau^* & \cosh \tau^* & 0 \\ 0 & 0 & 1 \end{bmatrix}$$
(7.21)

$$\tau = \nu z = \left(\kappa_y - \kappa_x - i\beta_o \left(n'_x - n'_y\right)\right) \frac{z'}{\cos \theta_o}$$
(7.22)

where $\kappa_{x,y}$ are the amplitude extinction coefficients of the volume for <u>x</u> and <u>y</u> polarisations. Note that if we cannot align the radar coordinates with the volume then the matrix term $\mathbf{R}(2\psi)$, which multiplies the whole matrix integral expression inside the brackets, causes a coherent mixing of terms that is difficult to interpret. We will show that the polarimetric optimizer automatically aligns the radar to the oriented volume. This result follows from knowledge of the explicit form of the matrix [K], which for this problem enables direct calculation of its eigenvalues and eigenvectors, and hence optimization parameters in closed form.

7.2.3 Optimum coherence values for oriented volume scattering

To account for the effects of wave propagation on the polarimetric response of an oriented volume, it is simpler to employ the covariance matrix [*C*] in the x/y basis rather than the coherency matrix [*T*]. Initially we set $\psi = 0$; that is, we assume that the radar and medium eigenpropagation coordinates are aligned. In this case we can explicitly invert the polarimetric covariance matrix as shown in equation (7.23):

$$C_{11} = \begin{bmatrix} c_{11}I_1 & 0 & c_{13}I_2 \\ 0 & c_{22}I_3 & 0 \\ c_{13}^*I_2^* & 0 & c_{33}I_4 \end{bmatrix} \Rightarrow C_{11}^{-1} = \frac{1}{f} \begin{bmatrix} c_{33}I_4 & 0 & -c_{13}I_2 \\ 0 & \frac{f}{c_{22}I_3} & 0 \\ -c_{13}^*I_2^* & 0 & c_{11}I_1 \end{bmatrix}$$
(7.23)

where $f = (c_{11}c_{33}I_1I_4 - c_{13}c_{13}^*I_2I_2^*) = (c_{11}c_{33} - c_{13}c_{13}^*)I_1I_4$, and similarly for the polarimetric interferometry matrix we can write the following factorization:

$$\Xi_{12} = U_{LP3}\Omega_{12}U_{LP3}^{-1} = e^{i\phi(z_o)} \begin{bmatrix} c_{11}I_5 & 0 & c_{13}I_6\\ 0 & c_{22}I_7 & 0\\ c_{13}^*I_8 & 0 & c_{33}I_9 \end{bmatrix}$$
(7.24)

The volume integrals $I_1 - I_9$ are defined in terms of the complex propagation constants $\beta_x = \beta_0 n'_x - i\kappa_x = \beta_0 n'_x - i\frac{\sigma_x}{2}$ and $\beta_y = \beta_0 n'_y - i\kappa_y = \beta_0 n'_y - i\frac{\sigma_y}{2}$ for

the two eigenpolarisations and β_z , the interferometric wavenumber, as follows:

$$I_{1} = \int_{0}^{h} e^{2\sigma_{x}z} dz \qquad I_{2} = \int_{0}^{h} e^{2i(\beta_{y}^{*} - \beta_{x})z} dz \qquad I_{3} = \int_{0}^{h} e^{2(\sigma_{x} + \sigma_{y})z} dz$$

$$I_{4} = \int_{0}^{h} e^{2\sigma_{y}z} dz \qquad I_{5} = \int_{0}^{h} e^{i\beta_{z}z} e^{2\sigma_{x}z} dz \qquad I_{6} = \int_{0}^{h} e^{i\beta_{z}z} e^{2i(\beta_{y}^{*} - \beta_{x})z} dz$$

$$I_{7} = \int_{0}^{h} e^{i\beta_{z}z} e^{(\sigma_{x} + \sigma_{y})z} dz \qquad I_{8} = \int_{0}^{h} e^{i\beta_{z}z} e^{-2i(\beta_{y}^{*} - \beta_{x})z} dz \qquad I_{9} = \int_{0}^{h} e^{i\beta_{z}z} e^{2\sigma_{y}z} dz$$

$$(7.25)$$

Hence the first part of the optimization matrix $[K_C]$ has the following form, which is diagonal if $I_4I_6 - I_2I_9 = I_8I_1 - I_2^*I_5 = 0$.

$$C_{11}^{-1}\Xi_{12} = \frac{e^{i\phi(z_0)}}{f} \begin{bmatrix} c_{33}I_4 & 0 & -c_{13}I_2 \\ 0 & \frac{f}{c_{22}I_3} & 0 \\ -c_{13}^*I_2^* & 0 & c_{11}I_1 \end{bmatrix} \begin{bmatrix} c_{11}I_5 & 0 & c_{13}I_6 \\ 0 & c_{22}I_7 & 0 \\ c_{13}^*I_8 & 0 & c_{33}I_9 \end{bmatrix}$$
(7.26)

From equation (7.25) we can easily show that both equations are satisfied for arbitrary medium parameters, as we have the following relationships:

$$I_4 I_6 = \int_0^h e^{2\sigma_y z} e^{i\beta_z z} e^{2i(\beta_y^* - \beta_x)z} dz = I_2 I_9$$

$$I_8 I_1 = \int_0^h e^{2\sigma_x z} e^{i\beta_z} e^{-2i(\beta_y^* - \beta_x)z} dz = I_2^* I_5$$
(7.27)

Hence the product $C_{11}^{-1} \Xi_{12}$ is diagonal. It follows that $K_c = C_{11}^{-1} \Xi_{12} C_{11}^{-1} \Xi_{12}^{*T}$ is also diagonal, which confirms that the optimum coherences are obtained when the radar coordinates are aligned with the medium axes. Furthermore, we can also find expressions for the complex diagonal values (the optimum coherences), as shown in equation (7.28):

$$\begin{split} \tilde{\gamma}_{1} &= \frac{(c_{11}c_{33}I_{4}I_{5} - c_{13}c_{13}^{*}I_{2}I_{8})}{(c_{11}c_{33} - c_{13}c_{13}^{*})I_{1}I_{4}} = \frac{I_{4}I_{5}}{I_{1}I_{4}} = f(\sigma_{x}) = \frac{2\sigma_{x}e^{i\phi(z_{0})}}{\cos\theta_{o}(e^{2\sigma_{x}h_{v}/\cos\theta_{o}} - 1)} \int_{0}^{h} e^{i\beta_{z}z'} e^{\frac{2\sigma_{x}z'}{\cos\theta_{0}}} dz' \\ \tilde{\gamma}_{2} &= \frac{I_{7}}{I_{3}} = f(\sigma_{x}, \sigma_{y}) = \frac{(\sigma_{x} + \sigma_{y})e^{i\phi(z_{0})}}{\cos\theta_{o}(e^{(\sigma_{x}\sigma_{y})h_{v}/\cos\theta_{o}} - 1)} \int_{0}^{h} e^{i\beta_{z}z'} e^{\frac{(\sigma_{x} + \sigma_{y})z'}{\cos\theta_{0}}} dz' \\ \tilde{\gamma}_{3} &= \frac{(c_{11}c_{33}I_{1}I_{9} - c_{13}c_{13}^{*}I_{6}I_{2}^{*})}{(c_{11}c_{33} - c_{13}c_{13}^{*})I_{1}I_{4}} = \frac{I_{1}I_{9}}{I_{1}I_{4}} = f(\sigma_{y}) = \frac{2\sigma_{y}e^{i\phi(z_{0})}}{\cos\theta_{o}(e^{2\sigma_{y}h_{v}/\cos\theta_{0}} - 1)} \int_{0}^{h} e^{i\beta_{z}z'} e^{\frac{2\sigma_{y}z'}{\cos\theta_{0}}} dz' \\ (7.28) \end{split}$$

By using the relationship between [*T*] and [*C*], the eigenvectors of K_c and $K_T = T_{11}^{-1} \Omega_{12} T_{11}^{-1} \Omega_{12}^{*T}$ are orthogonal, and given as shown in equation (7.29). Any mismatch between the radar and medium principal axis (the angle ψ) can

now be corrected by an inverse rotation of the eigenvectors of K_T , as shown in equation (7.30):

$$K_{T} \stackrel{eigenvectors}{\longleftrightarrow} \underbrace{w}_{1} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1\\ -1\\ 0 \end{bmatrix} \quad \underbrace{w}_{2} = \begin{bmatrix} 0\\ 0\\ 1 \end{bmatrix} \quad \underbrace{w}_{3} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1\\ 1\\ 0 \end{bmatrix}$$
$$K_{C} \stackrel{eigenvectors}{\longleftrightarrow} \underbrace{w}_{1} = \begin{bmatrix} 1\\ 0\\ 0 \end{bmatrix} \quad \underbrace{w}_{2} = \begin{bmatrix} 0\\ 0\\ 1 \end{bmatrix} \quad \underbrace{w}_{3} = \begin{bmatrix} 0\\ 1\\ 0 \end{bmatrix} \quad (7.29)$$
$$(K_{T}R(-\psi) \stackrel{eigenvectors}{\longleftrightarrow} \underbrace{w}_{1} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1\\ -\cos 2\psi\\ \sin 2\psi \end{bmatrix} \quad \underbrace{w}_{2} = \begin{bmatrix} 0\\ \sin 2\psi\\ \cos 2\psi \end{bmatrix} \quad \underbrace{w}_{3} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1\\ -\cos 2\psi\\ -\sin 2\psi \end{bmatrix} \quad (7.30)$$

 $R(\psi$

We see that the eigenvectors of [K] are orthogonal for oriented volume scattering, and also contain information about the orientation of the medium's eigenpolarisations, while the eigenvalues are related to the coherences for the corresponding eigenwave extinctions. As expected from physical arguments, the highest (lowest) coherence is obtained for the polarisation with the highest (lowest) extinction. The higher the extinction, the less penetration into the volume and hence the lower the effective volume decorrelation. This connection between the structure function and optimum coherences is summarized in Figure 7.6. Importantly, this provides our first example of an extended, nontrivial coherence loci. Figure 7.7 shows how the loci, defined by the three optimum coherence points, can be constructed. Note the following important features:

- 1) The three optima are rank ordered in coherence amplitude (radius inside the unit circle), with the highest coherence associated with the highest extinction propagation channel.
- 2) The three optima are also ranked in phase. If we take the phase of the bottom of the layer ($z = z_0$) as reference (shown as the black point





Fig. 7.6 Summary of physical interpretation of optimum coherence triplet for oriented volume scattering

Fig. 7.7 Coherence loci for oriented volume scattering: ideal case (left), and including SNR or temporal decorrelation (right)



Fig. 7.8 Coherence loci for an infinite half space with oriented volume scattering and varying extinctions

on the unit circle in Figure 7.7) then the lowest coherence amplitude is always closest in phase to this point, followed by the crosspolarised channel, and then the highest coherence point always has the highest phase shift.

3) The loci extension is caused by physical changes in volume decorrelation with polarisation, and so far we have ignored effects due to temporal and SNR decorrelation. These can be included in the analysis by allowing radial shifts, so extending the loci towards the origin, forming a new loci bounded by the dotted lines on the right side of Figure 7.7.

As an important special case we can consider the coherence region and loci for a semi-infinite oriented volume. Figure 7.8 shows the semicircular loci developed for the infinite random volume (equation (7.14)), and superimposed we show the three optima for the oriented volume. The coherence corresponding in Figure 7.6 to the polarisation with maximum extinction has the highest coherence and position closest to the top of the layer (which now corresponds to zero phase in this diagram). The minimum extinction lies furthest in phase, while the crosspolarised channel has a coherence intermediate in phase and amplitude between the maximum and minimum extinctions.

From a parameter estimation perspective the finite slab oriented volume (OV) model is interesting. We see that we have six observables (the phase and amplitude of coherence in three optimum polarisations), and yet we have only four unknowns (the layer depth, two values of extinction and phase of the bottom of the layer). This is a good starting point for developing robust algorithms for estimating these parameters from experimental data (see Chapter 8). However, there is one major problem to be addressed, in that we have so far ignored the presence of a 'hard' boundary behind the layer. In radar applications this is often a soil or rock surface beneath vegetation or snow/ice layer, and as we shall see, this considerably distorts the coherence loci shape. To consider such

issues we need to extend our approach to consider the coherence region for two-layer scattering problems.

7.3 The coherence loci for a two-layer scattering model

In this section we consider coherent scattering from a two-layer medium as shown schematically in Figure 7.9. A wave is incident at angle θ to the normal, and first impinges on the top layer, which we assume is a volume scatterer of depth $h_{\rm v}$. The bottom of this layer has position z_0 , defining the boundary between the two layers. Below this extends a second medium, with depth d. In what follows we shall assume that the mean dielectric constant of layer 2 is much greater than that of layer 1, and that $d >> \delta_p$, the penetration depth in layer 2. This has two important consequences for our analysis. Penetration into layer 2 is small compared to the depth of layer 1. In addition, we assume that the penetration into layer 2 is small enough to make the baseline scaled factor small; that is, $\beta_z \delta_p < 0.1$ (see equation (7.1)). In this case there is no significant volume decorrelation from layer 2. Although there may be negligible volume decorrelation from layer 2, the large contrast in mean dielectric constant across the boundary at $z = z_0$ implies that there will be a strong surface reflection. We conclude that the influence of the second layer is to act as a hard boundary behind the volume. As we shall see, however, this leads to significant complexity in the coherence region, mainly because the reflection and scattering from this boundary is polarisation sensitive, and also because of the complexities caused by multiple scattering, as we now consider.

Figure 7.10 shows a schematic representation of the four principal scattering mechanisms to be considered. On the left is shown the two principal direct mechanisms: volume backscattering from layer 1, and surface backscatter from the rough boundary at $z = z_0$. The first point to make is rather obvious from this diagram, but evidently the surface component is seen *through* layer 1, and hence the backscatter will depend not only on surface properties but also on two-way propagation through layer 1. Even in these simple mechanisms we see that the responses from the two layers are coupled in the final solution. This coupling effect becomes more apparent when we add multiple scattering mechanisms shown on the right-hand side of Figure 7.10. In the simplest case we can consider second-order interactions whereby backscatter can occur through two cascaded specular reflections—first from the surface, and then from elements in the scattering volume. Note that we actually have two scenarios to consider in such effects—the first running from A to B, and then the time-reversed path from B to A. It is the combination of these two mechanisms that maintains a





Fig. 7.9 Schematic representation of the geometry of a two-layer scattering problem





Fig. 7.11 Ray geometry explanation of phase centre location for dihedral scattering

symmetric backscatter matrix for reciprocal media (see Section 3.4.3). Importantly, even though these are second order, because the two can be specular (forward scattering with angle of incidence equal to angle of reflection), such second-order scattering contributions can be as large as, or larger than, the direct backscatter mechanisms themselves. Hence we cannot ignore them for a full development of the coherence loci for this two-layer problem. As we are considering coherent scattering, we must also concern ourselves with the phase of such second-order mechanisms.

Figure 7.11 shows a schematic representation of typical second-order scattering from a volume scattering element P at height h_p . We are concerned with the phase of the second-order signal from P (not the direct return) compared to R. This will depend on the range difference relative to a direct surface return coming from the surface at point R as shown. Point Q, being the specular reflection point on the surface, defines a triangle PQR as shown. It follows from the geometry of this triangle, shown enlarged on the right-hand side, that the distance QS + SP = 2SR. Combining this with recognition that PQ represents a wave front of the incident plane wave, which by definition is an equi-range contour, leads us to conclude that the range difference between P and R is zero for all heights h_p . This important result implies that in backscatter geometry, the phase difference between the second-order and direct surface scattering components is always zero.

There is, however, one further complication to be considered. If we now extend our analysis to consider across-track radar interferometry, then the second-order scattering behaves very differently for single and dual transmitter modes. In dual transmitter mode (which includes repeat-pass interferometry as a special case) we transmit and receive separately from each end of the base-line B. From each position the second-order scattering effects (being exactly in backscatter geometry) behave as shown in Figure 7.11, and consequently the phase difference across the baseline will be $\exp(i\beta_z z_0)$ —the same as for the direct surface scattering component from point R. Therefore, in dual transmitter mode the second-order scattering effects behave like an *effective* and additional surface component, with a phase corresponding to the underlying surface position and a coherence of unity, indicating zero volume decorrelation, even though the distributed volume is involved in the scattering mechanism (see equation (7.31)).

$$\tilde{\gamma}_{sv}^{2TX} = e^{i\phi(z_o)} \tag{7.31}$$

Note that this all follows from the special geometry of the triangle PQR in Figure 7.11. In terms of modifications of the structure function, we note that the second-order effects add an additional delta function contribution at $z = z_0$. Note also that from a polarimetric point of view, the second-order components have a polarisation signature (scattering $\alpha > \pi/4$) very different from direct surface scatter ($\alpha < \pi/4$). Hence, for dual transmitter modes we conclude that second-order surface–volume interactions and direct surface scattering are separable in polarimetry but not in interferometry.

Now consider the case of single transmit/dual receive interferometry. By definition this is a configuration that involves a small but non-zero bistatic scattering angle $\delta\theta$, as shown schematically in Figure 7.12. For the end of the baseline operating with both transmit and receive modes, the second-order



Fig. 7.12 Ray diagrams for bistatic dihedral scattering phase contributions

scattering will again have an exact backscatter geometry, and the effective phase centre lies on the surface at R (shown by the dash line in Figure 7.12). However, for the end of the baseline with a receive-only mode we obtain a height-dependent phase shift due to the small bistatic geometry. As the volume scattering elements are distributed over a range of heights from 0 to h_v , there will consequently be a volume decorrelation effect in this mode. To analyse further, consider separately the geometry of the surface–volume and volume– surface contributions shown in Figure 7.12. On the left we show the ray path for the surface–volume term. The scattering into a small bistatic angle $\delta\theta$ gives rise to a height-dependent phase given by equation (7.32). (Note the factor of 2π in place of 4π , because we are considering the single transmitter case.)

$$\phi_{SV}(h_p) = \frac{2\pi \sin \theta \delta \theta}{\lambda} h_p = \sin^2 \theta \beta_z h_p \quad \phi_{VS}(h_p) = -\frac{2\pi \sin \theta \delta \theta}{\lambda} h_p$$
$$= -\sin^2 \theta \beta_z h_p \quad (7.32)$$

Here β_z is the vertical wavenumber of the interferometer (assuming range spectral filtering has been employed), and h_p is the height of the volume scattering element (see Treuhaft, 1996, 2000a).

A similar argument applies to the volume-surface component shown on the right-hand side of Figure 7.12. However, this time we must consider the small bistatic angle $\delta\theta$ as arising from the surface specular point rather than the volume scatterer at P. In order to relate this to a height-dependent phase, we note that the surface scattering appears, from the baseline point of view, to come from a virtual point P_m which lies beneath the surface as shown in Figure 7.13, so that the distance PQ equals P_mQ. This has the effect of changing the sign of the interferometric phase as shown on the right in equation (7.32). These phase variations will combine and lead, via integration across the full depth of layer 1, to volume decorrelation. To calculate an expression for this decorrelation, we consider it as arising from an effective vertical profile function $f_2(z)$, which in general will have a shape different from that of the single scattering volume return $f_v(z)$. However, since the path length for second-order scattering through layer 1 is invariant to the height h_p of P (and equal to $2h_v/\cos\theta$), it follows that for an homogeneous layer (with an exponential profile $f_v(z)$, for example) the total extinction suffered by the wave is independent of the height h_p (that is, scattering elements at the top of layer 1 suffer propagation loss equal to those from the bottom of the layer). Under these circumstances $f_2(z)$ will be a



Fig. 7.13 Ray diagram for effective dihedral scattering contributions

uniform profile ranging from $-h_v$ to h_v from which we can calculate the decorrelation caused by second-order scattering interactions for single transmitter configurations as a SINC function, as shown in equation (7.33):

$$\begin{split} \tilde{\gamma}_{sv}^{1TX} &= e^{i\phi(z_o)} \frac{\int_0^{h_v} e^{i\sin^2\theta\beta_z z'} dz' + \int_{-h_v}^0 e^{i\sin^2\theta\beta_z z'} dz'}{2\int_0^{h_v} dz'} \\ &= e^{i\phi(z_o)} \frac{\int_{-h_v}^{h_v} e^{i\sin^2\theta\beta_z z'} dz'}{2h_v} \\ &= e^{i\phi(z_o)} \frac{\sin(\sin^2\theta\beta_z h_v)}{\sin^2\theta\beta_z h_v} \end{split}$$
(7.33)

Note that the mean phase centre still lies on the surface at $z = z_0$ (as it does for the dual transmitter case). However, we now have a radial shift towards the origin of the coherence diagram, with the amplitude of the coherence decreasing with increasing depth of layer 1.

We have seen in the above that second-order interactions cause some complexity in the analysis of coherent scattering from two-layer media. This is further compounded when we realize that in theory there is an infinite cascade of such higher-order surface-volume interactions to be considered. For example, on the far right of Figure 7.10 is a typical third-order interaction of surface-volume-surface scattering. Fortunately, for random media, such higher order interactions are usually very small compared to first and second order. Physically we can see that this arises because the backscatter level of such high order interactions is determined by a cascaded product of small quantities. The surface-volume-surface interaction, for example, involves the product of two surface reflections, which will be small, all multiplied by the backscatter rather than specular forward scatter from particles in the volume, which will generally be smaller. Add to this the increased effective propagation distance inside the lossy material of layer 1, and we can see, at least qualitatively, how interactions higher than second order can often in practice be ignored (although there are some notable exceptions such as scattering from complex man-made structures such as bridges and buildings). Nonetheless, for most remote sensing applications this result will allow us to justify the use of simpler second-order models in deriving the coherence loci for such problems.



Fig. 7.14 Ray diagram to locate phase centre for third-order scattering contribution

First, however, for completeness we consider in detail the coherence properties of third-order surface–volume–surface contributions. For simplicity we consider only the dual transmitter configuration. Figure 7.14 shows a schematic of the geometry concerned. Again we are interested, for interferometry, in how the effective range difference to a point P across the baseline varies as the height h_p is changed. In this case we see that the third-order effect has the same range variation as scattering from a virtual point P_m located a distance h_p beneath the surface, as shown. Therefore, even in the dual transmitter case we now obtain a height-dependent phase that will cause (complex) volume decorrelation and a loss of coherence amplitude combined with a (negative) phase bias.

We can calculate the level of this coherence by realizing that it has a corresponding structure function $f_3(z)$ that is extended below the surface into the range $0 < z < -h_v$. From this structure function we can then calculate the corresponding complex coherence from an integral of the general form shown



Fig. 7.15 Structure functions for various contributions to two-layer scattering

in equation (7.34). Note that the superscripted ^N represents the number of transmitters (1 or 2), and we have extended the range of the coherence integral from $-h_v$ to h_v to accommodate calculations for the virtual scattering points.

$$\tilde{\gamma}_{i}^{NTX} = e^{i\phi(z_{o})} \frac{\int_{-h_{v}}^{h_{v}} f_{i}^{NTX}(z') e^{i\beta_{z}z} dz'}{\left| \int_{-h_{v}}^{h_{v}} f_{i}^{NTX}(z') dz' \right|}$$
(7.34)

Figure 7.15 summarizes the form of the structure functions for various components of the two-layer scattering problem. In the first diagram we show the direct volume term, f_v^{2TX} , which has some arbitrary shape, bounded by the surface and the top height of the layer. Next we show the corresponding direct surface return, f_s^{2TX} , which is a simple delta function located on the surface. For the two-transmitter scenario this delta function also matches the second-order surface–volume scattering contribution, f_{sv}^{2TX} . However, the single-transmitter case has a uniform structure function extending across the full range, as shown in the lower diagram f_{sv}^{1TX} . Finally we show the structure function for thirdorder scattering f_{svs}^{2TX} . This has a totally negative extent, but again can lead to volume decorrelation with an arbitrary structure function bounded by the surface and top layer.

Of more general interest is the way in which these components combine to provide the overall coherence variation for a two-layer problem. To see this we need to incorporate all the mechanisms into a single generalized polarimetric interferometric formulation. The starting point is to define the coherent scattering vector \underline{k} as the sum of contributions at ends 1 and 2 of the baseline, as shown in equation (7.35). In these expressions, [P] is the vector propagation

matrix through layer 1 (see Section 4.2.6).

$$\begin{split} &\stackrel{}{\underbrace{\underline{k}}_{1}} = \left(\underline{k}_{v1} + [P_{s}]\underline{k}_{s1} + [P_{sv}]\underline{k}_{sv1} + [P_{svs}]\underline{k}_{svs1} + \cdots\right)}{\underline{k}_{sv2}} \\ \stackrel{}{\underbrace{\underline{k}}_{2}} = \left(\underline{k}_{v2} + [P_{s}]\underline{k}_{s2} + [P_{sv}]\underline{k}_{sv2} + [P_{svs}]\underline{k}_{svs2} + \cdots\right)} \\ &\stackrel{}{\Rightarrow} \begin{cases} T_{11} = \left\langle\underline{k}_{1}\underline{k}_{1}^{*T}\right\rangle = \left\langle\underline{k}_{v1}\underline{k}_{v1}^{*T}\right\rangle + [P_{s}]\left\langle\underline{k}_{s1}\underline{k}_{s1}^{*T}\right\rangle [P_{s}]^{*T} + [P_{sv}]\left\langle\underline{k}_{sv1}\underline{k}_{sv1}^{*T}\right\rangle [P_{sv}]^{*T} + \cdots \\ \Omega_{12} = \left\langle\underline{k}_{1}\underline{k}_{2}^{*T}\right\rangle = \left\langle\underline{k}_{v1}\underline{k}_{v2}^{*T}\right\rangle + [P_{s}]\left\langle\underline{k}_{s1}\underline{k}_{s2}^{*T}\right\rangle [P_{s}]^{*T} + [P_{sv}]\left\langle\underline{k}_{sv1}\underline{k}_{sv2}^{*T}\right\rangle [P_{sv}]^{*T} + \cdots \\ &\stackrel{}{\Rightarrow} \begin{cases} T_{11} = T_{V} + P_{s}T_{s}P_{s}^{*T} + P_{sv}T_{sv}P_{sv}^{*T} + \cdots \\ \Omega_{12} = \Omega_{v} + P_{s}\Omega_{s}P_{s}^{*T} + P_{sv}\Omega_{sv}P_{sv}^{*T} + \cdots \end{cases} \end{split}$$
(7.35)

We can then combine these vectors with averaging (which removes all cross-products between mechanisms—an expression of independent scattering mechanisms) to express the generalized polarimetry and interferometry as the sum of component matrices, as shown in equation (7.35). From these matrices we can then determine a general expression for the observed coherence as a function of polarisation, as shown in equation (7.36):

$$\tilde{\gamma}\left(\underline{w}\right) = \frac{\underline{w}^{*T}\Omega_{12}\underline{w}}{\underline{w}^{*T}T_{11}\underline{w}} = \frac{\underline{w}^{*T}(\Omega_{v} + P_{s}\Omega_{s}P_{s}^{*T} + P_{sv}\Omega_{sv}P_{sv}^{*T} + \cdots)\underline{w}}{\underline{w}^{*T}(T_{v} + P_{s}T_{s}P_{s}^{*T} + P_{sv}T_{sv}P_{sv}^{*T} + \cdots)\underline{w}}$$
$$= \frac{m_{0v}(\underline{w})\tilde{\gamma}_{v}(\underline{w}) + p_{s}(\underline{w})m_{0s}(\underline{w})\tilde{\gamma}_{s}(\underline{w}) + p_{sv}(\underline{w})m_{0sv}(\underline{w})\tilde{\gamma}_{sv}(\underline{w}) + \cdots}{m_{0v}(\underline{w}) + p_{s}(\underline{w})m_{0s}(\underline{w}) + p_{sv}(\underline{w})m_{0sv}(\underline{w}) + \cdots}$$
(7.36)

Here we have rewritten each term as a product of three components: its normalized radar cross section m_0 , its coherence contribution $\tilde{\gamma}$, and a propagation factor *p* that attenuates each contribution according to the propagation paths involved. Note that we can write this expression as the product of total radar backscatter times total observed coherence, as shown in equation (7.37):

$$\tilde{\gamma}(\underline{w})m_0(\underline{w}) = m_{0\nu}(\underline{w})\tilde{\gamma}_{\nu}(\underline{w}) + p_s(\underline{w})m_{0s}(\underline{w})\tilde{\gamma}_s(\underline{w}) + p_{s\nu}(\underline{w})m_{0s\nu}(\underline{w})\tilde{\gamma}_{s\nu}(\underline{w}) + \cdots$$
(7.37)

This gives us a procedure for deriving the coherence loci for two-layer problems, by which we first need to calculate the three elements for each mechanism, and then combine them as shown in equation (7.36). As mentioned earlier, for lossy layers this series converges quickly, and we need not consider the complexity of scattering higher than second order. To illustrate this we now develop three particular forms of this model that are widely used in the literature: a coherent two-layer version of the water cloud model (IWCM), the closely related random-volume-over-ground or RVOG model, and the oriented-volume-over-ground or OVOG model.

7.4 Important special cases: RVOG, IWCM and OVOG

An important class of models can be generated by making the following assumptions about the two-layer scattering problem:

- Assume dual transmitter operation only (including repeat-pass as a special case), so removing the coherence loss due to surface-volume interactions. In this case the direct surface and surface-volume multiple scattering contributions both have structure functions given by a Dirac delta function.
- 2. Assume an exponential structure function for the direct volume return. This amounts to the physical assumption of a layer of uniform density characterized by a mean wave extinction σ_e , which may nonetheless be a function of polarisation.
- 3. Assume that the layer is lossy enough and the surface rough enough that third- and higher-order interactions can be ignored.

By allowing polarisation dependence of extinction we are essentially assuming that layer 1 is an oriented volume, and this leads to the most complicated form of such two–layer scenarios: the oriented-volume-over-ground (OVOG) model. Before considering this case, however, we first develop a pair of models based on the simpler assumption that the propagation is scalar and does not depend on polarisation.

By assuming a random volume for layer 1, it follows that the propagation factors simplify, as they become independent of polarisation and are a function only of a single mean extinction coefficient σ_e . There are two important models that make use of this approach: the random-volume-over-ground (RVOG) model (Treuhaft, 1996; Papathanassiou, 2001; Cloude, 2003), and the interferometric water cloud model (IWCM) model (Askne, 2003, 2007). They differ primarily in their assumptions about the importance of temporal decorrelation. In RVOG it is common to assume that $\gamma_t = \gamma_{snr} = 1$, which indicates a dominance of volume decorrelation over all other sources; while for IWCM it is commonly assumed that γ_t is dominant. RVOG is therefore better suited to single-pass or low-frequency large spatial/low temporal baseline repeat pass interferometry, while IWCM has been applied mainly to high-frequency small spatial/large temporal baseline applications. We now examine the polarisation dependence of each of these models, with a view to deriving their coherence loci.

7.4.1 The random-volume-over-ground (RVOG) model

In the RVOG approach the structure function for the two-layer problem reduces to the simple form shown in Figure 7.16. Note that the second-order (dihedral) scattering effects are included as a coherent addition to the direct surface return. Importantly, the polarisation dependence of coherence is now restricted to a single term, as we now demonstrate. The RVOG model leads to a coherence as shown in equation (7.38):

$$\tilde{\gamma}(\underline{w})m_{0}(\underline{w}) = m_{0\nu}(\underline{w})\tilde{\gamma}_{\nu} + p_{s}m_{0s}(\underline{w})e^{i\phi(z_{o})} + p_{s\nu}m_{0s\nu}(\underline{w})e^{i\phi(z_{o})}$$

$$\Rightarrow \tilde{\gamma}(\underline{w}) = e^{i\phi(z_{o})}\frac{m_{0\nu}(\underline{w})\tilde{\gamma}_{\nu}e^{-i\phi(z_{o})} + p_{s}m_{0s}(\underline{w}) + p_{s\nu}m_{0s\nu}(\underline{w})}{m_{0\nu}(\underline{w}) + p_{s}m_{0s}(\underline{w}) + p_{s\nu}m_{0s\nu}(\underline{w})}$$

$$\Rightarrow \tilde{\gamma}(\underline{w}) = e^{i\phi(z_{o})}\frac{\tilde{\gamma}_{\nu o} + \mu(\underline{w})}{1 + \mu(w)}$$
(7.38)

where μ is the ratio of effective surface-to-volume scattering. We also note that the volume decorrelation component does not depend on polarisation, and is



Fig. 7.16 Composite structure function for RVOG model

given explicitly as shown in equation (7.39):

$$\tilde{\gamma}_{\nu}e^{-i\phi(z_{o})} = \tilde{\gamma}_{\nu o} = \frac{p_{1}(e^{p_{2}h_{\nu}} - 1)}{p_{2}(e^{p_{1}h_{\nu}} - 1)} \begin{cases} p_{1} = \frac{2\sigma_{e}}{\cos\theta_{o}} \\ p_{2} = \frac{2\sigma_{e}}{\cos\theta_{o}} + i\beta_{z} \end{cases}$$
(7.39)

Significantly, only the parameter μ changes with polarisation. This is the ratio of *effective* (sum of direct and second-order) surface-to-volume scattering. We can develop an explicit form for μ as shown in equation (7.40):

$$\mu\left(\underline{w}\right) = \frac{p_s m_{0s}(\underline{w}) + p_{sv} m_{0sv}(\underline{w})}{m_{0v}(\underline{w})}$$
(7.40)

This can be further simplified by realizing that the propagation factors for direct surface and second-order interactions are the same and given by equation (7.41):

$$p_s = p_{sv} = e^{-2\sigma_e h_v \sec\theta} \tag{7.41}$$

Moreover, we can express the volume scattering contribution in the denominator of equation (7.40) as a function of the scalar extinction, layer depth and angle of incidence (all independent of polarisation), and a polarisation-dependent scattering cross-section, as shown in equation (7.42):

$$m_{ov}\left(\underline{w}\right) = \frac{\cos\theta}{2\sigma_e} \left(1 - e^{-2\sigma_e h_v \sec\theta}\right) m_v\left(\underline{w}\right)$$
(7.42)

Here m_v has a corresponding diagonal coherency matrix, as shown in equation (7.43), where *s* depends on the mean particle shape in the volume (s = 0 for spheres 0.5 for prolate spheroids).

$$m_{\nu}(\underline{w}) = \underline{w}^{*T} [T_{\nu}] \underline{w} = m_{HH+VV} \underline{w}^{*T} \begin{bmatrix} 1 & 0 & 0 \\ 0 & s & 0 \\ 0 & 0 & s \end{bmatrix} \underline{w} \quad 0 \le s \le 0.5 \quad (7.43)$$

From this we see that the cross-section can vary in the range $m_{HH+VV} \leq m_v(\underline{w}) \leq sm_{HH+VV}$. For a dipole cloud, for example, we note there is only a 3-dB variation of RCS with polarisation. The factor μ , however, can have a much wider dynamic range than this, as shown in equation (7.44):

$$\mu\left(\underline{w}\right) = \frac{2\sigma_e}{m_v(\underline{w})} \frac{(m_{0s}(\underline{w}) + m_{0sv}(\underline{w}))e^{-2\sigma_e h_v \sec \theta}}{\cos \theta (1 - e^{-2\sigma_e h_v \sec \theta})}$$
$$= \frac{2\sigma_e}{\cos \theta (e^{2\sigma_e h_v \sec \theta} - 1)} \frac{m_{0s}(\underline{w}) + m_{0sv}(\underline{w})}{m_v(\underline{w})}$$
(7.44)

Here we see that the numerator depends directly on the variation of 'effective' surface scattering with polarisation. We can assume that this has reflection symmetry and a corresponding variation with polarisation, as shown in equation (7.45):

$$m_{0s}(\underline{w}) + m_{0sv}(\underline{w}) = \underline{w}^{*T}([T_s] + [T_{sv}])\underline{w}$$

= $\underline{w}^{*T}[T_{es}]\underline{w}$
= $m_{HH+VV}\underline{w}^{*T}\begin{bmatrix} 1 & t_{12} & 0\\ t_{12}^* & t_{22} & 0\\ 0 & 0 & t_{33} \end{bmatrix}\underline{w}$ (7.45)

Here the subscripted _{es} denotes the effective surface components. The polarisations that maximize and minimize the μ ratio will be of interest in establishing the coherence loci for RVOG. To find these we need to solve the following eigenvalue equation arising from optimization of the μ ratio, as shown in equation (7.46) (see Section 4.2.2.2):

$$\max \frac{\underline{w}^{*T}[T_{es}]\underline{w}}{\underline{w}^{*T}[T_{v}]\underline{w}} \Rightarrow [T_{v}]^{-1}[T_{es}]\underline{w}_{opt} = \lambda \underline{w}_{opt}$$
(7.46)

Explicitly, we then obtain the following optimum ratio values as a function of the mean particle shape and normalized effective surface coherency matrix elements.

$$\frac{m_s^{hh+vv}}{m_v^{hh+vv}} \begin{bmatrix} 1 & 0 & 0 \\ 0 & s & 0 \\ 0 & 0 & s \end{bmatrix}^{-1} \begin{bmatrix} 1 & t_{12} & 0 \\ t_{12}^* & t_{22} & 0 \\ 0 & 0 & t_{33} \end{bmatrix} \\
\Rightarrow \begin{cases}
\lambda_1 = \frac{m_s^{hh+vv}}{2m_v^{hh+vv}} \max\left(1 + \frac{t_{22}}{s} \pm \sqrt{\left(1 - \frac{t_{22}}{s}\right)^2 + \frac{4|t_{12}|^2}{s}}\right) \\
\lambda_2 = \frac{m_s^{hh+vv}}{2m_v^{hh+vv}} \min\left(1 + \frac{t_{22}}{s} \pm \sqrt{\left(1 - \frac{t_{22}}{s}\right)^2 + \frac{4|t_{12}|^2}{s}}\right) \\
\lambda_3 = \frac{m_s^{hh+vv}}{m_v^{hh+vv}} \frac{t_{33}}{s}
\end{cases}$$
(7.47)

We note two important features of this solution:

- 1) Firstly, the optimum eigenvectors (of which there are three from equation (7.46)) are not orthogonal (since $T_v^{-1} T_{es}$ is neither symmetric nor Hermitian). Contrast this with the case of polarimetric interferometry for an oriented volume, which yields a set of three orthogonal scattering mechanisms (see equation (7.29)). This is often an important signature of the presence of multilayer scattering effects in polarimetric interferometry.
- 2) The second important point to note is that the optimum μ values are given by the eigenvalues λ_1, λ_2 and λ_3 . The ratio λ_1/λ_3 then gives a measure of the maximum dynamic range of μ with polarisation. We shall see below that this also impacts on the size of the coherence loci for the RVOG model. Note that these optima will not in general occur for a fixed polarisation basis (the Pauli basis, for example), as the structure of $[T_{es}]$ will change with surface conditions (roughness, moisture, and so on). Therefore, if we wish to make use of these optimum values for

parameter estimation, for example, then we need to employ a more adaptive processing approach, based on coherence optimization, to be able to exploit these extreme values. We are now in a position to determine the coherence loci for the RVOG model as follows.

7.4.2 Polarisation coherence loci for RVOG

The shape of the coherence loci for the RVOG model is best developed by first rewriting the expression for RVOG coherence (equation (7.38)) in the following form:

$$\tilde{\gamma}(\underline{w}) = e^{i\phi(z_o)} \left(\tilde{\gamma}_{vo} + \frac{\mu(\underline{w})}{1 + \mu(\underline{w})} (1 - \tilde{\gamma}_{vo}) \right)$$
$$= e^{i\phi(z_o)} (\tilde{\gamma}_{vo} + F(\underline{w}) (1 - \tilde{\gamma}_{vo}))$$
(7.48)

Here we have deliberately isolated the polarisation dependence in a single term, $F(\underline{w})$. This factor is real non-negative, and lies in the range $0 \le F(\underline{w}) \le 1$, with limits occurring at one end for pure volume scattering ($\mu = 0$), and at the other by pure surface scattering ($\mu = \infty$). Hence F(w) is directly the fraction of (effective) surface scattering in the observed signal. With $\tilde{\gamma}_{vo}$ a fixed complex number, independent of polarisation, this, then, is the equation of a straight line in the complex plane, going through the points $\tilde{\gamma}_{vo}$ and $e^{i\phi(z_o)}$, as shown in Figure 7.17. The coherence loci for the RVOG model is therefore a straight line.

Note that in practice not all of this line will be visible from experimental data, and it is here that the dynamic range of μ becomes important. In reality there will only ever be visible some limited segment of this line, corresponding to the variation of *F* from μ_{min} to μ_{max} . Note importantly, however, that this line is not radial, as the volume coherence is always complex, and thus there is a phase as well as an amplitude variation with polarisation. Note also that the variation of





Fig. 7.18 Variation of coherence magnitude in RVOG model for various phase angles

coherence amplitude with increasing μ is not monotonic. As μ increases from zero the coherence passes through a local minimum, the position of which can be calculated exactly from the coherence expression in equation (7.48), as shown in equation (7.49). The value of μ that produces the minimum coherence (point of closest approach of the line to the origin of the coherence diagram) is given by $\mu = 0$, except when the parameter F_{\min} in equation (7.49) is greater than zero. This is a function not only of the volume coherence amplitude, but also its phase. Figure 7.18 shows an example for volume coherence of 0.8 with 30-degree steps in phase from 0° to 180°, at which point the line passes through the origin and hence the minimum coherence is zero.

$$\begin{split} \tilde{\varphi}(\underline{w}) &= e^{i\phi(z_o)}(\tilde{\gamma}_{vo} + F(\underline{w})(1 - \tilde{\gamma}_{vo})) \\ \Rightarrow L &= \left|\tilde{\gamma}(\underline{w})\right|^2 = a + bF + cF^2 \\ \Rightarrow \frac{dL}{dF} &= b + 2cF \Rightarrow F_{\min} = -\frac{b}{2c} = \frac{\left|\tilde{\gamma}_{vo}\right|^2 - \operatorname{Re}(\tilde{\gamma}_{vo})}{(1 - \tilde{\gamma}_{vo})\left(1 - \tilde{\gamma}_{vo}^*\right)} \\ \Rightarrow \left|\tilde{\gamma}(\underline{w})_{\min}\right| &= \left|\tilde{\gamma}_{vo} + \frac{\left|\tilde{\gamma}_{vo}\right|^2 - \operatorname{Re}(\tilde{\gamma}_{vo})}{(1 - \tilde{\gamma}_{vo})\left(1 - \tilde{\gamma}_{vo}^*\right)}(1 - \tilde{\gamma}_{vo})\right| = \left|\frac{\operatorname{Im}(\tilde{\gamma}_{vo})}{(1 - \tilde{\gamma}_{vo}^*)}\right| \\ \Rightarrow if \ F_{\min} > 0 \Rightarrow \mu_{\min} = 10 \log_{10}\left(\frac{\left|\tilde{\gamma}_{vo}\right|^2 - \operatorname{Re}(\tilde{\gamma}_{vo})}{1 - \operatorname{Re}(\tilde{\gamma}_{vo})}\right) \end{split}$$
(7.49)

We see that for small volume phase shifts ϕ the minimum is given by small $\mu < -30$ dB, but as the phase increases so that $\text{Re}(\tilde{\gamma}_{vo}) > |\tilde{\gamma}_{vo}|^2$ or $\phi > \cos^{-1}(|\tilde{\gamma}_{vo}|)$, then the minimum coherence occurs for higher values of μ . Figure 7.19 shows how the μ for minimum coherence changes for the example shown in Figure 7.18. Note that for phase angles up to $\cos^{-1}(0.8) = 37^{\circ}$ the minimum


Fig. 7.19 μ required for minimum coherence in RVOG model versus phase

(

coherence is given by the volume only ($\mu = 0$) point. However, for phase shifts above this the minimum occurs for a mixture of surface and volume scattering. For high phase shifts the minimum occurs for an almost equal mixture of surface and volume scattering.

However, we have seen that the RVOG assumes an exponential structure function in the volume, and hence the coherence amplitude and phase are not independent quantities. In fact the phase centre for the volume-only component of RVOG must lie between halfway and the top of layer 1 (see Figure 5.18). In other words, for RVOG it follows that the phase of the interferogram for the volume-only channel $\phi \geq \frac{\beta_z h_v}{2}$. Coupled to this is the realization that for RVOG the coherence amplitude can never be less than the zero extinction limit; that is, $|\tilde{\gamma}_{vol}| \geq \text{sinc}(\frac{\beta_z h_v}{2})$.

This then begs the question as to whether the minimum coherence for the RVOG model can *ever* be given by $\mu = 0$. For this to be possible the following inequality must apply:

$$\phi_{\min} < \cos^{-1}(|\tilde{\gamma}_{\min}|) \Rightarrow \frac{\beta_z h_v}{2} < \cos^{-1}\left(\operatorname{sinc}\left(\frac{\beta_z h_v}{2}\right)\right) \quad 0 \le \frac{\beta_z h_v}{2} \le \pi$$
(7.50)

However, this inequality is never satisfied for the RVOG model, and hence we conclude that for RVOG the minimum coherence is never given by the $\mu = 0$ volume scattering channel. There is always some mixture of surface and volume scattering that combines to produce a minimum. This rather surprising result follows from the assumed form of the structure function (an exponential). This exposes a weakness of the RVOG model: that its assumption of a uniform layer with a simple extinction profile ignores any variations due to vertical structure in volume scattering. We now turn to consider, in more detail, how structure variations are dealt with in RVOG.

7.4.3 Structural ambiguity in RVOG

Before leaving the RVOG model we first consider one important extension: its generalization to arbitrary volume structure functions. By maintaining the assumption of a random volume in layer 1 and polarisation-dependent delta function contributions from the effective surface components, we can generalize the RVOG approach to arbitrary structure functions, as shown schematically in Figure 7.20. Importantly, this modification maintains the line as coherence loci, but it changes the relationship between the phase and coherence amplitude of the volume only ($\mu = 0$) point. This we call a structured-volume-over-ground (SVOG) model, which has the general form shown in equation (7.51):

$$\begin{split} \tilde{\gamma}(\underline{w}) &= e^{i\phi(z_o)} \left(\tilde{\gamma}_{vo} + \frac{\mu(\underline{w})}{1 + \mu(\underline{w})} \left(1 - \tilde{\gamma}_{vo} \right) \right) \\ \tilde{\gamma}_{vo} &= e^{i\beta_z z_o} \frac{\int_0^{h_v} f_v(z') e^{i\beta_z z'} dz'}{\int_0^{h_v} f_v(z') dz'} \\ &= e^{i\beta_z z_0} e^{ik_v} \frac{(1 + a_0)f_0 + a_1 f_1 + a_2 f_2 + \dots + a_n f_n}{(1 + a_0)} \end{split}$$
(7.51)



Fig. 7.20 Structure function for general structured volume-over-ground or SVOG model

For this more general SVOG model the relationship between coherence amplitude and phase is not as restrictive as it is for RVOG, and so the minimum coherence can indeed be the volume-only coherence, under the general condition that the phase $\phi < \cos^{-1}(|\tilde{\gamma}_{vo}|)$.

The classical RVOG model can be made to accommodate changes in structure by varying the extinction over a sufficiently wide range. However, this gives rise to a structural ambiguity in RVOG, in that we can fit RVOG to non-RVOG situations simply by adjusting the model parameters. A simple example of this ambiguity is shown in Figure 7.21. Here we show a simple case of vertical structure: a layering of the volume, with scattering from a thin top layer, and surface reflection from a position separated from this volume by a gap.

The coherence model for this three-layer problem can be written as shown in equation (7.52). This still has a linear coherence loci, but the coherence amplitude of the volume-only channel will be small, while its phase will be



Fig. 7.21 Vertical structural ambiguity in the RVOG model

large. It is possible to fit RVOG to this structure (as shown in the upper portion of Figure 7.21). However, to explain the combination of high coherence and large phase offset we need to employ an effective extinction for the medium that is much larger than the actual value.

$$\tilde{\gamma}(\underline{w}) = e^{i\beta_{z}z_{o}} \left(e^{i\beta_{z}h_{c}} \tilde{\gamma}_{vo} + \frac{\mu(\underline{w})}{1+\mu(\underline{w})} \left(1 - e^{i\beta_{z}h_{c}} \tilde{\gamma}_{vo} \right) \right)$$

$$\tilde{\gamma}_{vo} = \frac{\int_{h_{c}}^{h_{v}} e^{\frac{2\alpha_{c}h_{v}}{\cos\theta}} e^{i\beta_{z}z'} dz'}{\int_{h}^{h_{v}} e^{\frac{2\alpha_{c}h_{v}}{\cos\theta}} dz'}$$
(7.52)

7.4.4 The coherence loci for IWCM

In this section we consider a model closely related to RVOG, but one that arose independently out of generalizations of the water cloud model (WCM; see Section 3.5.1) (Askne, 1997, 2003, 2007). This model, called the interferometric water cloud model (IWCM), places more emphasis on the temporal changes in volume and surface scattering, and was developed for representing the coherence observed by repeat-pass, high-frequency, small spatial baseline radar systems. The model shares the exponential structure function for volume scattering and assumed random volume for layer 1 of RVOG (although in its most general form it further splits the effective extinction into wave extinction in the canopy and the fraction of gaps between the vegetation; see Askne (2003)). However, it explicitly includes a vertical temporal stability function, as shown schematically in Figure 7.22. This function is 1 where the volume is stable, and 0 where unstable. While arbitrary stability functions (and their corresponding Fourier-Legendre expansions) can be envisaged, usually the simplest assumption of a uniform pulse function (zeroth-order Legendre function) with amplitude γ_{tv} for the volume component and γ_{ts} and γ_{tsv} for the surface elements, is taken. Equation (7.53) shows the general form of coherence for this model:

$$\tilde{\gamma}(\underline{w})m_{0}(\underline{w}) = m_{0\nu}(\underline{w})\gamma_{t\nu}\tilde{\gamma}_{\nu o}e^{i\phi(z_{o})} + p_{s}m_{0s}(\underline{w})\gamma_{ts}e^{i\phi(z_{o})} + p_{s\nu}m_{0s\nu}(\underline{w})\gamma_{ts\nu}e^{i\phi(z_{o})}$$
(7.53)

There are two forms of this model that deserve special attention. In the first we consider its form for short spatial/long temporal baselines and high radar frequency, where volume decorrelation is very small and temporal effects for both surface and volume dominate. We shall call this the high-frequency or HF-IWCM. In the second form—more closely linked to RVOG—we consider longer spatial/shorter temporal baselines and low radar frequency when temporal effects are mixed with significant levels of volume decorrelation. This we call the low-frequency or LF-IWCM.

In the first case the β_z value is small and the mean wave extinction is high (because of the high-frequency approximation), and the volume decorrelation can therefore be approximated by a unitary phase shift, as shown in



Fig. 7.22 Composite structure function for the IWCM model

equation (7.54):

$$\tilde{\gamma}_{\nu o} = \frac{p_1(e^{p_2 h_\nu} - 1)}{p_2(e^{p_1 h_\nu} - 1)} \begin{cases} p_1 = \frac{2\sigma_e}{\cos \theta_o} \\ p_2 = \frac{2\sigma_e}{\cos \theta_o} + i\beta_z \end{cases}$$
$$\stackrel{\lim \sigma_e \to \infty}{\longrightarrow} \tilde{\gamma}_{\nu o} \approx \frac{p_1}{p_2} e^{i\beta_z h_\nu} = \frac{1}{1 + i\frac{\beta_z \cos \theta_o}{2\sigma_e}} e^{i\beta_z h_\nu} \qquad (7.54)$$
$$\approx e^{-i\frac{\beta_z \cos \theta_o}{2\sigma_e}} e^{i\beta_z h_\nu} = e^{i\beta_z \left(h_\nu - \frac{\cos \theta_o}{2\sigma_e}\right)}$$

With this in place the HF-IWCM takes the following form:

$$\tilde{\gamma}(\underline{w})m_{0}(\underline{w}) = \tilde{\gamma}_{tv}\frac{m_{v}(\underline{w})\cos\theta}{2\sigma_{e}}\left(1 - e^{-2\sigma_{e}h_{v}\sec\theta}\right)e^{i\beta_{z}(h_{v} - \frac{\cos\theta_{o}}{2\sigma_{e}})} + e^{-2\sigma_{e}h_{v}\sec\theta}\tilde{\gamma}_{es}m_{0es}(\underline{w})$$
(7.55)

or in terms of coherence it can be written thus:

$$\tilde{\gamma}(\underline{w}) = \frac{\tilde{\gamma}_{tv} \frac{m_v(\underline{w}) \cos\theta}{2\sigma_e} \left(1 - e^{-2\sigma_e h_v \sec\theta}\right) e^{i\beta_z(h_v - \frac{\cos\theta_o}{2\sigma_e})} + e^{-2\sigma_e h_v \sec\theta} \tilde{\gamma}_{es} m_{0es}(\underline{w})}{\frac{m_v(\underline{w}) \cos\theta}{2\sigma_e} \left(1 - e^{-2\sigma_e h_v \sec\theta}\right) + e^{-2\sigma_e h_v \sec\theta} m_{0es}(\underline{w})}{e^{i\beta_z(h_v - \frac{\cos\theta_o}{2\sigma_e})} + \tilde{\gamma}_{es}\mu(\underline{w})}$$

$$= \frac{\tilde{\gamma}_{tv} e^{i\beta_z(h_v - \frac{\cos\theta_o}{2\sigma_e})} + \tilde{\gamma}_{es}\mu(\underline{w})}{1 + \mu(\underline{w})}$$
(7.56)

Again we see that the parameter μ —the surface-to-volume scattering ratio—is important in determining the coherence loci for this model. The loci in this case is a triangle, with two vertices on the unit circle and one at the origin ($\gamma_{tv} = \gamma_{tes} = 0$), as shown in Figure 7.23. An important simplified case of HF-IWCM arises in the limit $\gamma_{tv} = 0$ and $\gamma_{es} = 1$. This occurs in vegetation problems, for example, when wind-driven temporal change destroys the coherence completely from layer 1, while the underlying surface scattering contributions show no change. In this case we move along the radial line OP in Figure 7.23, for which the coherence loci depends entirely on μ , as shown in equation (7.57):

$$\tilde{\gamma}(\underline{w}) = e^{i\phi(z_o)} \frac{\mu(\underline{w})}{1 + \mu(\underline{w})}$$
(7.57)

Note that this has same form of coherence variation with polarisation as SNR decorrelation, as demonstrated in Figure 7.24 (compare this with Figure 5.12).

In the second form of this model—the LF-IWCM—we consider the limit of larger spatial/lower temporal baselines and low radar frequency combined with volume-only temporal decorrelation γ_{tv} due, for example, to short-term wind-blown effects in vegetation cover. This model is closely related to RVOG, since the volume coherence can no longer be approximated by a phase shift.



Fig. 7.23 Coherence loci for the IWCM model

Fig. 7.24 Coherence variation with surfaceto-volume ratio (μ) for the IWCM model

The two models—RVOG and LF-IWCM—can be connected by rewriting the latter in the following form:

$$\tilde{\gamma}(\underline{w}) = e^{i\phi(z_o)}(\gamma_{tv}\tilde{\gamma}_{vo} + F(\underline{w})(1 - \gamma_{tv}\tilde{\gamma}_{vo}))$$
(7.58)

Here we see that the coherence loci remains a non-radial line segment, but that the fixed point representing the volume is shifted towards the origin by the scale



Fig. 7.25 Coherence loci for the RVOG model with temporal decorrelation

factor γ_{tv} . This amounts to a rotation and stretch of the coherence line about the surface topography point, as shown in Figure 7.25 (Papathanassiou, 2003).

In conclusion, we have shown that the coherence loci for a two-layer random volume over ground scattering problem is a line segment in the complex plane. This line is radial when temporal effects dominate, and shifts to a non-radial phase variant line as volume decorrelation becomes more important. In both cases we note two important features. The first is the importance of the ratio μ , being the ratio of effective surface-to-volume scattering. The second key point is that the line passes through the unit circle at the surface phase point. We shall see later that this provides us with a method for correcting for vegetation bias in radar interferometry by line fitting and by finding this intersection.

7.4.5 The coherence loci for OVOG

In the previous section we considered the case when layer 1 is a random volume, and showed that the coherence loci is then a straight line in the complex plane. A natural extension of this approach is to consider layer 1 as an oriented volume. In this case the polarimetry becomes more complex, as discussed in Section 7.2.2. However, we shall see that the coherence loci may still be obtained as a simple extension of the RVOG approach.

The OVOG model maintains the assumption of a uniform layer with an exponential structure function, but is characterized by a pair of eigenpolarisations for propagation through the medium. These orthogonal states then define a triplet of structure functions, as shown schematically in Figure 7.26. The states with highest extinction XX and lowest extinction YY are separated by the crosspolarised channel XY. The effective surface components (shown as a line at $z = z_0$) are viewed through the polarisation filter of volume 1, which distorts their apparent polarimetry (see Section 4.2.6). In the presence of combined surface and volume scattering we must now consider a triplet of coherence formulae—one



Fig. 7.26 Composite structure function for the OVOG model

for each eigenpolarisation combination, as shown in equation (7.59):

$$\begin{split} \tilde{\gamma}_{xx} &= e^{i\phi} \frac{\gamma_{vo} \left(2\sigma_{x}, h_{v}\right) + \mu_{xx}}{1 + \mu_{xx}} = e^{i\phi} (\tilde{\gamma}_{vo}^{xx} + F_{xx}(1 - \tilde{\gamma}_{vo}^{xx})) \\ \tilde{\gamma}_{xy} &= e^{i\phi} \frac{\gamma_{vo} \left(\sigma_{x} + \sigma_{y}, h_{v}\right) + \mu_{xy}}{1 + \mu_{xy}} = e^{i\phi} (\tilde{\gamma}_{vo}^{xy} + F_{xy}(1 - \tilde{\gamma}_{vo}^{xy})) \quad (7.59) \\ \tilde{\gamma}_{yy} &= e^{i\phi} \frac{\gamma_{vo} \left(2\sigma_{y}, h_{v}\right) + \mu_{yy}}{1 + \mu_{yy}} = e^{i\phi} (\tilde{\gamma}_{vo}^{yy} + F_{yy}(1 - \tilde{\gamma}_{vo}^{yy})) \end{split}$$

Here both volume and surface scattering have polarimetric coherency matrices with reflection rather than azimuthal symmetry, and the surface-to-volume scattering ratios include the effects of propagation distortion. The dynamic range of μ can be developed using a modification of the procedure used in equation (7.46), as shown in equation (7.60):

$$\mu_{opt} \to \max \frac{\underline{w}^{*T}[P][T_{es}][P]^{*T}\underline{w}}{\underline{w}^{*T}[T_{v}]\underline{w}} \Rightarrow [T_{v}]^{-1}[P][T_{es}][P]^{*T}\underline{w}_{opt} = \lambda \underline{w}_{opt}$$
(7.60)

Where [P] is a propagation distortion matrix (see Section 4.2.6). While it is now not so easy to develop an analytic solution for the eigenvalues of this optimization, we can obtain an estimate of the coherence loci by using a simple geometrical argument, as follows.

Each term in the triplet of coherences in equation (7.59) has the same form as the RVOG model, and thus corresponds to a line segment in the complex plane. Furthermore, as these eigenstates bound the oriented volume solution (see Section 7.2.2) it follows that the loci must be contained within the triplet of lines defined in equation (7.59). Figure 7.27 shows the resulting triangular coherence loci for the OVOG model. The coherence for each polarisation



Fig. 7.27 Coherence loci for the OVOG model

is constrained to move up and down its own straight line inside the unit circle, depending on the μ ratio. The three lines for the eigenstates define the boundary of this region, coming to a focus at the ground topography point ϕ , and having a spread ψ , as shown in Figure 7.27. Importantly, this spread depends on the differential extinction in the volume layer and not on the μ values or surface topography. In the special case that $\psi = 0$ we again obtain the random-volume-over-ground RVOG model. We also note that the OVOG model requires that the crosspolarised XY coherence line lies between the XX and YY lines.

7.4.6 The oriented-volume-under-ground (OVUG) model

Finally, we consider an important variation of the OVOG model, applicable to cases where scattering occurs from the top surface of layer 1 and at the same time h_v tends to infinity, so that we can effectively ignore the influence of scattering from the layer 1–2 interface. Such a model can be used for analysis of thick layers, as occur, for example, in high-frequency land-ice applications, where scattering from the top air-ice interface usually dominates that from the bottom ice/rock interface. This scenario is summarized in Figure 7.28, in which is shown the corresponding structure function. The coherence function for this problem can be derived in a similar manner to the OVOG model, and is shown in equation (7.61). The key difference here is the phase of the volume term, which now lies below the surface reference rather than above it (equation (7.62)).

$$\tilde{\gamma}(\underline{w}) = e^{i\phi(z_o)}(\tilde{\gamma}^*_{\infty}(\underline{w}) + F(\underline{w})(1 - \tilde{\gamma}^*_{\infty}(\underline{w})))$$
(7.61)

$$\tilde{\gamma}_{\infty}^{*}(\underline{w}) = \frac{1}{1 + i\frac{\beta_{z}\cos\theta}{2\sigma_{e}(w)}}$$
(7.62)

The coherence loci for this problem can be obtained as an extension of the OV region, as shown in Figure 7.29. Here again we see a region formed by three bounding lines for the eigenpolarisations emanating from the surface point, with variations along each line given by the fraction of surface-to-volume scattering, $F(\underline{w})$. Again the RVUG or random volume version of this model is obtained as a limiting case when the extinctions become equal and we obtain the single line coherence region, as shown on the right-hand side of Figure 7.29.





Fig. 7.28 Composite structure function for the OVUG model

Fig. 7.29 Coherence loci for the OVUG model (left) and OVOG model (right)

8

Parameter estimation using polarimetric interferometry



Fig. 8.1 Geometry of two-layer scattering problem

In the previous chapter we developed the form of the backscatter polarimetric interferometric coherence loci for a two-layer scattering model. We now turn to consider algorithms for the inverse problem for such a case; that is, we consider methods for estimation of parameters of the two-layer model from observations of the coherence variation with polarisation (Papathanassiou, 2001; Cloude, 2000b, 2000c, 2003; Stebler, 2002; Ballester-Berman, 2005; Praks 2007). We start by identifying the key parameters of interest by reference to the schematic diagram shown in Figure 8.1. Based on this we can identify the following important parameters of interest in remote sensing:

- 1. The position of the bottom of the layer, z_0 (or its associated interferometric phase, ϕ_0). This if often called the underlying or true surface topography or ground position in vegetation and snow/ice applications.
- 2. The depth of layer 1, h_v , which may correspond to the height of vegetation or depth of a snow layer, depending on the application.
- 3. In some applications (such as land-ice penetration), interest centres on the position (phase) of the top of layer 1, especially when its depth tends to infinity (Dall, 2003; Sharma, 2007). In this case the main application is to compensate the penetration depth into the layer so as to locate the true surface position.
- 4. The structure function in layer 1, f(z). For exponential models such as RVOG and OVOG this amounts to estimation of a pair of extinction coefficients, while in more general terms it amounts to estimation of the Fourier–Legendre spectrum of the structure function for the layer.
- 5. The surface-to-volume scattering ratio, μ . This function, when combined with the total backscatter cross-section, can be used to separate scattering contributions from layers 1 and 2 and hence isolate volume or surface scattering for further study (Cloude, 2004, 2005a).

We now turn to consider estimation techniques for each of these in turn.

8.1 Surface topography estimation

There are three basic approaches to the estimation of underlying surface topography (Papathanassiou, 2001; Sagues, 2000; Cloude, 2000c, 2003). The simplest is an extension of conventional interferometry, employing the phase of an interferogram for some selected surface-dominated polarisation vector \underline{w}_{s} . The second approach is to employ two polarisation states to remove phase bias from the top layer. Finally, we can use multiple polarisations and least squares correction to phase bias. We now consider each of these in turn.

In the simplest case the phase of the surface component can be estimated directly from the coherence, as shown in equation (8.1):

$$\hat{\phi} = \arg(\tilde{\gamma}_{\underline{w}_s}) \quad \mathbf{0} \le \hat{\phi} < 2\pi$$
(8.1)

By subsequently employing phase unwrapping, the surface topography can then be estimated. The precision of this estimate (given by the height variance) depends on baseline and the coherence amplitude of the interferogram, as shown in equation (8.2):

$$\sigma_{h} \approx \frac{R_{0} \sin\theta}{B_{\perp}} \frac{\lambda}{4\pi} \sigma_{\phi} \quad \sigma_{\phi} \leq \sqrt{\frac{1 - \left|\gamma_{\underline{w}_{s}}\right|^{2}}{2L \left|\gamma_{\underline{w}_{s}}\right|^{2}}}$$
(8.2)

where the Cramer–Rao bound (minimum value) of the phase variance σ_{ϕ} for a given number of looks *L* is given on the right-hand side of equation (8.2) (Seymour, 1994). This, of course, resorts to conventional interferometry in the limiting case of bare surfaces ($h_v = 0$), but in other cases is made complicated by the phenomenon of phase bias.

We have seen that the volume coherence for layer 1 is complex and hence contributes a phase offset from the surface itself, and so equation (8.1) will generally overestimate the surface position for RVOG and OVOG, and underestimate it for OVUG. Hence it is clear that polarisation \underline{w}_s should be chosen so as to minimize this bias and optimize the accuracy of the estimate. A second objective must be to choose \underline{w}_s to also maximize the SNR, so as to minimize the decorrelation due to noise and hence optimize the precision. In RVOG, OVOG and OVUG these requirements amount to maximization of μ , the surface-tovolume ratio. The best polarisation to use would therefore be that given by equation (7.47) or (7.60). However, as we have no *a priori* knowledge of the separate volume and surface component coherency matrices, we cannot make direct use of this equation. Instead we must employ an indirect solution, as we now investigate.

The problem is that there is no single polarisation that always maximizes μ . For a bare surface ($h_v = 0$) at low frequencies when the Bragg surface scattering model is valid, a good choice is VV, as HH has less scattered power and thus lower SNR, and HV is zero. A better choice still is HH+VV, as the zero polarimetric phase difference leads to an even better SNR than VV. Clearly, the optimum would weight the Bragg scattering matrix elements to maximize the SNR. For higher frequencies and rougher surfaces the depolarisation increases, the difference between HH and VV becomes less, and polarisation plays less of a role in bare surface parameter estimation. In general, therefore, an unbiased coherence optimizer would provide a suitably adaptive solution. For bare surfaces the constrained optimizer of equation (6.22) would provide a good choice (as shown again in equation (8.3)). Note, however, that in order to implement such an optimizer we require access to full scattering matrix data, so as to be able to estimate the component matrices $\hat{T}_{11}, \hat{T}_{22}, \hat{\Omega}_{12}$.

$$[\hat{T}]^{-1}[\hat{\Omega}_{H}]\underline{w} = \lambda(\phi)\underline{w} \qquad \begin{cases} [\hat{\Omega}_{H}] = \frac{1}{2}\left(\hat{\Omega}_{12}e^{i\phi} + \hat{\Omega}_{12}^{*T}e^{-i\phi}\right) \\ [\hat{T}] = \frac{1}{2}\left(\hat{T}_{11} + \hat{T}_{22}\right) \end{cases}$$

$$\xrightarrow{\max[\lambda(\phi)]} \underline{w}_{opt} \Rightarrow \gamma_{opt} = \frac{\underline{w}_{opt}^{*T}[\hat{\Omega}_{H}]\underline{w}_{opt}}{\underline{w}_{opt}^{*T}[\hat{T}]\underline{w}_{opt}}$$

$$(8.3)$$

The above analysis breaks down, however, in case layer 1 has non-negligible thickness. We can set a suitable threshold on the product of wavenumber and layer thickness to estimate this breakpoint, such as $\beta_z h_v < 0.1$ (see equation 7.1), for the surface approximation to hold. If the product exceeds this threshold then we require a different strategy to optimize estimation of surface topography as follows.

When the layer thickness can no longer be ignored, we face complications arising not only from phase bias and increased volume decorrelation, but a change in scattering mechanism. This arises especially when the dihedral second-order scattering is dominant. In this case HH is often preferred to VV (the opposite of the bare surface case), as it has a higher specular reflection coefficient at the surface, and VV will in this case have a lower μ . By the same reasoning, the Pauli choice HH–VV is sometimes selected in preference to HH+VV, as this has an even higher μ than HH, due to the 180-degree polarimetric phase shift that occurs in the case of a dominant second-order scattering scenario.

A second problem also arises in the case that layer 1 is a random volume, when it acts to depolarise the scattered wave with a high entropy. This implies that the volume scattering 'contaminates' every polarisation vector \underline{w} , and consequently that it is impossible to find a candidate \underline{w}_s which contains surface-only scattering. This means that the phase bias due to volume scattering in layer 1 will be present across the whole of polarisation space. The best we can do is again try to select the \underline{w}_s with maximum μ . However, there is no longer a guarantee that the coherence amplitude optimizer of equation (8.3) will correspond to the maximum μ (see the discussion in Section 7.4.2). While the optimizer will still guarantee the highest coherence and hence the highest precision, it no longer guarantees the highest accuracy, because of the presence of phase bias. To proceed further we need to consider methods for the removal of this bias.

8.1.1 Phase bias removal

We can make use of the SVOG (with RVOG as a special case) model to remove the phase bias and improve the accuracy of surface topography estimation. We have seen that the SVOG model predicts linear coherence loci in the complex plane. Importantly, this line intersects the unit circle at the desired surface topography point $e^{i\phi}$. Therefore, if we start by selecting an arbitrary polarisation \underline{w}_1 and evaluate its interferometric coherence $\tilde{\gamma}_1$ it will lie somewhere on this line, generally displaced from the desired unit circle point by some unknown phase bias. However, if we now choose a second polarisation state \underline{w}_2 , and the only condition we set on \underline{w}_2 is that it have a higher surface-to-volume scattering ratio than the first, so that $\mu_2 > \mu_1$, then it follows that we can find the unit circle point from a line fit as follows. The idea is to use $\tilde{\gamma}_1$ as a fixed point on the line and relate $\tilde{\gamma}_2$ by a scale factor F_2 along the line towards the unit circle. In this way the two coherences can be related as shown in equation (8.4). We see that the desired ground topography point is embedded in these equations, and we can solve for it directly.

$$\frac{\tilde{\gamma}_1}{\tilde{\gamma}_2 = \tilde{\gamma}_1 + F_2(e^{i\phi} - \tilde{\gamma}_1)} \right\} \Rightarrow e^{i\phi_o} = \frac{\tilde{\gamma}_2 - \tilde{\gamma}_1(1 - F_2)}{F_2} \quad 0 \le F_2 \le 1 \quad (8.4)$$

Here we see that the phase term is obtained as a weighted average of the two complex coherences. If F_2 tends to unity then it represents the desired surface point, and $\tilde{\gamma}_2$ is taken as the solution. However, in general there will be phase bias present in both channels (because of depolarisation in layer 1), and hence this mixture formula is required to compensate for this bias. There remains a problem in that to solve for the surface topography we first require an estimate of the factor F_2 . This can be obtained directly from the estimated coherences $\hat{\gamma}_1$ and $\hat{\gamma}_2$ by forming the product $e^{i\phi}e^{-i\phi} = 1$, using equation (8.4) to obtain a quadratic. Taking the root that makes F positive, we obtain the solution shown in equation 8.5:

$$\hat{\phi} = \arg(\tilde{\gamma}_{2} - \tilde{\gamma}_{1} (1 - F_{2})) \quad \mathbf{0} \le F_{2} \le 1$$

$$AF_{2}^{2} + BF_{2} + C = 0 \Rightarrow F_{2} = \frac{-B - \sqrt{B^{2} - 4AC}}{2A}$$

$$A = |\tilde{\gamma}_{1}|^{2} - 1 \quad B = 2\operatorname{Re}((\tilde{\gamma}_{2} - \tilde{\gamma}_{1}).\tilde{\gamma}_{1}^{*}) \quad C = |\tilde{\gamma}_{2} - \tilde{\gamma}_{1}|^{2}$$
(8.5)

Note that unlike the simple phase algorithm (equation (8.1)), this requires coherence estimates in both amplitude and phase, and hence is susceptible to non-compensated errors in coherence such as those due to SNR or temporal effects. Also, of course, the estimates of coherence themselves have some variance due to their stochastic nature and the finite number of looks *L* used in the estimator (the coherence region). For stability of the phase estimate we therefore need to ensure that F_2 is as large as possible. (If F_2 tends to zero so that the two points are close together, then large errors can result.) Some care is therefore required in the selection of \underline{w}_1 and \underline{w}_2 . There are three strategies used in making an appropriate selection:

In physics-based selection we use our understanding of surface and volume scattering to select the two channels. For example, the low-frequency Bragg surface model predicts zero or (for higher-order forms of the model) very low levels of crosspolarisation HV from a flat surface. On the other hand, volume scattering from a cloud of anisotropic particles can yield high levels of HV (for a dipole cloud only 3 dB below the maximum RCS). Hence HV is often chosen as a candidate for the \underline{w}_1 channel. The \underline{w}_2 channel can likewise be selected on the assumption that specular second-order scattering is dominant, and so HH or HH–VV are good choices as they are likely to satisfy the requirement that $\mu_2 > \mu_1$. In summary, the direct physics-based approach produces allocations of the two channels such as those shown by the two examples in equation (8.6):

$$\tilde{\gamma}_{1} = \tilde{\gamma}_{HV} \quad \tilde{\gamma}_{2} = \tilde{\gamma}_{HH-VV}$$
or
$$\tilde{\gamma}_{1} = \tilde{\gamma}_{HV} \quad \tilde{\gamma}_{2} = \tilde{\gamma}_{HH}$$
(8.6)

Note that the lower option is particularly well suited to dual polarised active systems that can transmit only linear H polarisation but receive H and V components (see Chapter 9).

If more specific information is available about the scattering problem to hand, based, for example, on direct EM scattering model simulations, then these assignments can of course be modified as appropriate. Although such selections may match very well a specific application or dataset, they are generally not sufficiently robust for widespread application. For this reason we turn to a second approach, based instead on phase optimization, that adapts itself to variations in the data.

8.1.2 Coherence separation optimization

We have seen that phase bias can be removed under the assumptions of the SVOG model by fitting a line between two coherence values. Furthermore, best results will be obtained if we employ two polarisation states with the *maximum* difference in μ , as these will be less sensitive to fluctuation noise in the coherence region estimates for a fixed number of looks *L*. Under the SVOG model, μ impacts directly on the phase centre of the interferogram, so that as μ decreases so the phase bias increases. This indicates that the optimum pairing of polarisation vectors \underline{w}_1 and \underline{w}_2 to choose would be those corresponding to the ends of the linear coherence loci in the complex plane. One way to estimate these is to employ coherence *separation* rather than coherence amplitude optimization (see Section 6.2). In this method we employ fully polarimetric data to calculate the following eigenvalue problem (see Section 6.2.3).

$$[T]^{-1}[\Omega_{H}(\phi)]\underline{w} = \lambda (\phi) \underline{w} \xrightarrow{\substack{\phi \\ \phi}} \Delta \tilde{\gamma}_{opt}$$

$$\tilde{\gamma}_{1} = \frac{\underline{w}_{a}^{*T}[\Omega_{H}]\underline{w}_{a}}{\underline{w}_{a}^{*T}[T]\underline{w}_{a}}$$

$$\tilde{\gamma}_{2} = \frac{\underline{w}_{b}^{*T}[\Omega_{H}]\underline{w}_{b}}{\underline{w}_{b}^{*T}[T]\underline{w}_{b}}$$

$$(8.7)$$

For each phase angle ϕ we find the distance between the maximum and minimum eigenvalues. By finding the maximum of this distance as a function of ϕ we then automatically align the solution with the axis of the linear coherence region and use these as an estimate of the coherence loci bounds. We then find the polarisation scattering mechanisms \underline{w}_a and \underline{w}_b from the corresponding eigenvectors. From these we can derive the two coherence $\tilde{\gamma}_{max}$ and $\tilde{\gamma}_{min}$ for use in the topographic phase estimation algorithm of equation (8.5).

However, we face a potential problem with all these line-fitting ideas to ensure that we always choose the correct rank ordering of the two coherences, remembering that we must ensure that $\mu_2 > \mu_1$ to find the correct topography point. In fact this is a general problem with all phase bias removal algorithms based on the assumption of a linear coherence loci. By definition, a line intersects the unit circle at two points (see Figure 8.2), one of which is the true topographic phase, while the other represents a false solution obtained for the line fit technique by exchanging the rank order of coherences. There are several ways to resolve this





rank-ordering dichotomy, and two common techniques use physical arguments based on scattering theory or a comparison of the interferometric bias levels of the two solutions. We now consider both of these.

In the physical approach we again employ our expectations for the nature of polarimetric scattering in the channel with the highest μ value. For example, the surface-dominated channel should have a scattering vector close to the form \underline{w}_s shown in equation (8.8), where α is less than $\pi/4$ for direct surface scattering, and greater than $\pi/4$ for dihedral second-order scattering.

$$\underline{w}_{s} = \begin{bmatrix} \cos \alpha & \sin \alpha e^{i\varphi} & 0 \end{bmatrix}^{T} \\ \underline{w}_{v} = \begin{bmatrix} 0 & 0 & 1 \end{bmatrix}^{T} \end{cases} \xrightarrow{\mu_{s} > \mu_{v}} \begin{cases} \tilde{\gamma}_{1} = \tilde{\gamma}_{\min} & \text{if } |\underline{w}_{\max}^{*T} \underline{w}_{v}| < |\underline{w}_{\min}^{*T} \underline{w}_{v}| \\ \tilde{\gamma}_{1} = \tilde{\gamma}_{\max} & \text{if } |\underline{w}_{\max}^{*T} \underline{w}_{v}| > |\underline{w}_{\min}^{*T} \underline{w}_{v}| \end{cases}$$

$$(8.8)$$

Similarly, the orthogonal state \underline{w}_v should match the volume scattering (and have a lower μ). In this way we can develop an algorithm for assigning the optimum phase states in the correct rank order for the line fit algorithm, as shown on the right-hand side of equation (8.8).

In the second interferometric approach we decide on rank ordering by employing the phase difference between the calculated unit circle intersection point and assumed low μ coherence. Knowing that layer 1 is above (or below) the surface allows us to calculate these phase differences using the same clockwise (or anticlockwise) rotation around the coherence diagram. If we repeat this for both rank permutations, then one of the phase shifts will be much larger than the other and can be rejected (especially if it is known that the layer depth is less than the π height of the interferometer). If we define ϕ_{max} as the unit circle phase estimate obtained when we propose $\tilde{\gamma}_{max}$ as the high μ channel estimate, and likewise ϕ_{min} when we propose $\tilde{\gamma}_{max}$, then we can decide on the most likely rank ordering as follows:

$$\Delta_{\max} = \arg(\tilde{\gamma}_{\min}e^{-i\hat{\phi}_{\max}}) \\ \Delta_{\min} = \arg(\tilde{\gamma}_{\max}e^{-i\hat{\phi}_{\min}}) \end{cases} \rightarrow \begin{cases} \tilde{\gamma}_1 = \tilde{\gamma}_{\min} & \text{if } \Delta_{\max} < \Delta_{\min} \\ \tilde{\gamma}_1 = \tilde{\gamma}_{\max} & \text{if } \Delta_{\max} > \Delta_{\min} \end{cases}$$
(8.9)

So, for example, in Figure 8.2, if we assume that layer 1 is above the surface for anticlockwise phase rotation, then $\Delta_{\text{max}} = 270^{\circ}$ and $\Delta_{\text{min}} = 90^{\circ}$, and therefore according to equation (8.9) we would select $\tilde{\gamma}_1 = \tilde{\gamma}_{\text{max}}$ on the assumption that the layer thickness should be less than the π height of the interferometer.

8.1.3 Total least squares (TLS) surface topography estimation

So far we have considered line fit techniques that employ only two coherence values, selected either on the basis of scattering physics or by employing coherence optimization. A more robust version of this approach is to employ multiple polarisation channels (N > 2) and use a least squares line fit to the multiple complex data points. In this way we can avoid problems of any selected pair of points becoming too close, and thus minimize errors in the surface topography estimation.

We start by using the linear coherence loci assumption to generate a linear relationship between the real and imaginary parts of coherence, as shown in Figure 8.3. The problem then reduces to estimation of the two coefficients M and C. In ordinary least squares (LS) estimation we would find the M and C that minimize the sum of squares of the vertical distance between line and data points, Δy , as shown in Figure 8.3. However, this assumes that noise is only found in one coordinate, whereas for coherence estimation both real and imaginary parts are subject to statistical fluctuations (see equation (5.34)). A better approach, therefore, is to employ a total least squares solution (TLS) that accounts for errors in both x and y. Geometrically, the TLS approach amounts to using a different measure of distance: the perpendicular distance R_i , at right in Figure 8.3, and related to Δy as shown in equation (8.10):

$$R_i = \Delta y \cos \theta = \frac{\Delta y}{\sqrt{1 + \tan^2 \theta}} = \frac{\Delta y}{\sqrt{1 + M^2}}$$
(8.10)



Fig. 8.3 Total least squares line fit

If we then make the simplest assumption that the unknown fluctuation errors are the same in x and y (the modifications, if they are not, are straightforward, but complicate the notation), the function to be minimized now has the following form:

$$\sum_{i} R_{i}^{2} = \frac{1}{1 + M^{2}} \sum_{i} (y_{i} - C - Mx_{i})^{2}$$
(8.11)

This differs from the conventional LS approach only in the pre-multiplier, which is itself a function of M. We can then find the stationary points of this function by differentiation, to yield the following:

$$0 = \frac{\partial R}{\partial C} = \frac{1}{1+M^2} \sum_{i} -2(y_i - C - Mx_i)$$

$$0 = \frac{\partial R}{\partial M} = \frac{1}{1+M^2} \sum_{i} -2(y_i - C - Mx_i)x_i$$

$$-\frac{2M}{(1+M^2)^2} \sum_{i} (y_i - C - Mx_i)^2$$
(8.12)

From the first term we obtain a direct solution for the estimate of C as follows:

$$\hat{C} = \frac{1}{N} \left(\sum_{i} y_i - \hat{M} \sum_{i} x_i \right) = \bar{y} - \hat{M}\bar{x}$$
(8.13)

Then, by substituting the first equation in the second and collecting terms we obtain an estimate of M as the root of a quadratic, as shown in equation (8.14):

$$c_2 \hat{M}^2 + c_1 \hat{M} + c_0 = 0 \Rightarrow \hat{M} = \frac{-c_1 \pm \sqrt{c_1^2 - 4c_2 c_0}}{2c_2}$$
 (8.14)

where the three coefficients are defined in terms of the data points as follows:

$$c_{0} = -\sum_{i} (x_{i} - \bar{x})(y_{i} - \bar{y})$$

$$c_{1} = \sum_{i} \left\{ (x_{i} - \bar{x})^{2} - (y_{i} - \bar{y})^{2} \right\}$$

$$c_{2} = \sum_{i} (x_{i} - \bar{x})(y_{i} - \bar{y})$$
(8.15)

This then provides us with a method for fitting a line to an arbitrary number of polarisation channels. We can then use the estimates of M and C to find the two unit circle intersection points as shown in Figure 8.4. These two points can



Fig. 8.4 Example of total least squares line fit to complex coherence data

be found explicitly in terms of M and C as shown in equation (8.16):

$$\begin{cases} x^{2} + y^{2} = 1\\ y = \hat{M}x + \hat{C} \end{cases} \Rightarrow \begin{cases} x_{p} = \frac{-\hat{M}\hat{C} \pm \sqrt{\hat{M}^{2} - \hat{C}^{2} + 1}}{1 + \hat{M}^{2}}\\ y_{p} = \hat{M}x_{p} + \hat{C} \end{cases} \Rightarrow e^{i\phi} = x_{p} + iy_{p} \end{cases}$$
(8.16)

Clearly, this TLS approach will suffer from the same rank-ordering dichotomy encountered in Section 8.1.2. Indeed, in the TLS case this problem is arguably more serious, as there is no easy way to employ the physical selection process described in equation (8.8). As there are multiple polarisation channels, and not just two being used, it is difficult to decide which are surface- or volume-dominated. For this reason the TLS approach is often combined with the phase approach of equation (8.9) to decide which phase point to use as the topography estimate.

Finally, we note that some care is required in the choice of N polarisations to ensure that there is sufficient diversity of μ along the line segment in the complex plane. For this reason, typical selections involve the three Pauli channels (HH+VV, HH–VV and HV), augmented by the linear channels HH and VV, as well as the three optimum states from either constrained or unconstrained algorithms to produce a sample set in the region N = 8–12 for estimation.

8.1.4 OVOG: surface topography with differential extinction

In the previous section we outlined various algorithms for the removal of vegetation bias and consequent estimation of true surface phase in polarimetric interferometry. The main assumption behind these techniques was that of a linear coherence loci, which we have seen implies that layer 1 scatters with azimuthal symmetry. While this may be a valid assumption for many applications, it is violated for an important class of problems when layer 1 displays reflection scattering symmetry and behaves as an oriented volume. In this case we have seen that the coherence loci is formed by a fan of three lines, emanating from the unit circle topography point (see Figure 7.27). In theory, therefore, we cannot strictly apply the above algorithms to the OVOG model. However, there are two important classes of OVOG applications that deserve special mention (Ballester-Berman, 2005, 2007; Lopez-Sanchez, 2007).

The first involves applications with only weak differential extinction, in which case the fan angle ψ (see Figure 7.27) is small and the OVOG region approaches a straight line, or at the other extreme, high differential extinction combined with small minimum extinction. This latter scenario is very important in that it can lead to a wide dynamic range in μ , with the low extinction channel dominated by surface scattering and the high extinction by volume scattering. In this case the line fit approach of TLS or phase optimization still provides a good, if approximate, solution to surface topography estimation, as shown in Figure 8.5. However, in all OVOG cases it must be realized that there is an additional source of error due to the separation of volume terms in the complex plane, and this can lead to large errors if extra care is not taken to make full use of the μ spectrum for the problem. In this case, therefore, it is good to use either the TLS with a wide diversity of polarisations, or the coherence separation optimization technique to ensure that the maximum and minimum μ values are being fully exploited.

Another class of OV problems of interest are for layers of effective infinite depth, when we can assume that $\mu = 0$ in *all* polarisation channels. In this limit we essentially obtain the oriented volume or OV problem as a limiting case of OVOG. We have seen that the coherence loci for this problem is a semicircle (Figure 7.8), and we can devise a simple algorithm for top surface phase estimation based on a circle fit to the data, as follows. The first point to note is that the circle must intersect the unit circle and the origin, and hence has



Fig. 8.5 Line fit for topography estimation under the OVOG model



Fig. 8.6 Topography estimation for the OV model

a fixed radius of 0.25. The second point is that the topographic phase is simply related to the coordinates of the centre C of the circle as shown in Figure 8.6. By combining these two observations we can set a linear least squares formulation for the two unknown coordinates of C as shown in equation (8.17):

$$(p - p_o)^2 + (q - q_o)^2 = r^2$$

$$r^2 = p_o^2 + q_o^2 = 0.25$$

$$\Rightarrow 2 \begin{bmatrix} p_{xx} & q_{xx} \\ p_{xy} & q_{xy} \\ p_{yy} & q_{yy} \end{bmatrix} \cdot \begin{bmatrix} p_o \\ q_o \end{bmatrix} = \begin{bmatrix} p_{xx}^2 + q_{xx}^2 \\ p_{xy}^2 + q_{xy}^2 \\ p_{yy}^2 + q_{yy}^2 \end{bmatrix}$$

$$\Rightarrow [A]\underline{x} = \underline{b}$$

$$\Rightarrow \hat{x} = ([A]^T [A])^{-1} [A]^T \underline{b}$$
(8.17)

Here we make use of the real and imaginary parts of the coherence in the three eigenpolarisation combinations XX, XY and YY to fit the best-constrained circle to the data. This then allows us to estimate the top surface position, even though we are assuming there is no scattering from this interface and only volume scattering is occurring. This is useful when either the top surface is very smooth or there is a small dielectric contrast between free space and layer 1. Note that this technique does not work if $\mu > 0$ in any channel, as this has the effect of pulling the coherences off the circle and towards the topography point. In this case we must resort to the approximation used in Figure 8.5.

We have seen that there are several possibilities for using *a priori* assumptions about the coherence loci for the two-layer problem, to devise algorithms for estimation of the surface topography and hence to effect phase bias removal. We now turn to consider a similar approach to the estimation of a second important physical parameter: the height of the top layer.

8.2 Estimation of height $h_{\rm v}$

In this section we consider algorithms for the estimation of the top layer height h_v using single baseline polarimetric interferometry (Cloude, 2001b, 2001c, 2003; Papathanassiou, 2001, 2005; Stebler, 2002; Yamada, 2001; Praks, 2007). The approach will be to assume particular forms for the coherence loci for the two–layer problem, exploit knowledge of the topographic phase from the previous section, and use various scattering models to obtain an estimate of h_v from complex coherence.

One of the simplest approaches to this problem is to use the phase difference between interferograms as a direct estimate of layer depth (Cloude, 1998). In general terms we then estimate the coherence in two polarisation channels: \underline{w}_v , which is volume scattering only and has a phase centre near the top of the layer; and \underline{w}_s , which is surface dominated and has a phase centre near the surface. By forming the phase difference between these interferograms and scaling by the interferometric wavenumber β_z we obtain the following estimation algorithm:

$$\hat{h}_{\nu} = \frac{\arg(\tilde{\gamma}_{\underline{w}_{\nu}}\tilde{\gamma}_{\underline{w}_{s}}^{*})}{\beta_{z}}, \quad \beta_{z} = \frac{4\pi\,\Delta\theta}{\lambda\sin\theta}$$
(8.18)

Here again the arg(..) function is defined in the range 0 to 2π . Although this is a simple algorithm to implement it has some severe drawbacks in that the layer depth estimate so obtained is generally underestimated. The problems stem from the difficulty in finding polarisations with phase centres at the top and bottom of the layer. We have seen that because of depolarisation in layer 1 there will be some volume scattering present in all polarisation channels, and so the phase centre of \underline{w}_s , for example, will always be located above the true surface (due to phase bias as discussed in Section 5.2.4). Likewise, we have seen that the phase centre of the volume scattering component can lie anywhere between halfway up and the top of layer 1, only reaching the top in case of infinite extinction in the RVOG model, or more generally a vertical structure function which is a delta function at $z = z_0 + h_v$.

The phase bias issue for the surface channel can be compensated somewhat by using our estimate for true surface topography, so that equation (8.18) takes the modified form shown in equation (8.19):

$$\hat{h}_{\nu} = \frac{\arg(\tilde{\gamma}_{\underline{w}_{\nu}}e^{-i\phi})}{\beta_{z}}$$
(8.19)

Note that $\hat{\phi}$ can be estimated either from the data itself or from some external source such as a reference digital surface model (DSM). We can further improve this algorithm by matching it to the optimization process used in equation (8.7). In particular we can make use of the optimum coherence furthest in phase from the surface topography point as the ideal \underline{w}_v channel. This then acts to maximize the height of the phase centre of \underline{w}_v in layer 1. However, there still remains the problem of compensating the volume scattering channel for variations in structure function f(z). For example, if the structure function is uniform then the phase optimum will still only reach halfway up layer 1, and so the phase estimate of equation (8.19) will be only one half the true layer depth.

To try to resolve this, we note that $\tilde{\gamma}_{\underline{W}_{\nu}}$ is complex and so has two degrees of freedom, of amplitude as well as phase. However, we have so far made use

only of the phase information. The idea is therefore to try to use the coherence amplitude of $\tilde{\gamma}_{\underline{w}_{\nu}}$ to help compensate for variations in the structure function to obtain a better estimate of h_{ν} . We shall see that there are various ways of doing this, but a good starting point is to use a Fourier–Legendre expansion of the structure function f(z) in terms of an infinite series with coefficients a_{io} , which are then related to the coherence as shown in equation (8.20) (see Section 5.2.4 for a derivation of the functions f_i).

$$\tilde{\gamma} = e^{i\phi}e^{ik_{\nu}}(f_0 + a_{10}f_1 + a_{20}f_2 + ...) \quad a_{i0} = \frac{a_i}{1 + a_0}, k_{\nu} = \frac{\beta_z h_{\nu}}{2}$$
(8.20)

8.2.1 First-order inverse coherence model

In order to use the infinite series of equation (8.20) with experimental data we must first truncate the series at some finite order. The simplest non-trivial case is to truncate at first order, as shown in equation (8.21):

$$\tilde{\gamma} = e^{ik_v} e^{i\phi_o} (f_0 + a_{10}f_1 + R_1) \approx e^{i(k_v + \phi_0)} \left(\frac{\sin k_v}{k_v} + ia_{10} \left(\frac{\sin k_v}{k_v^2} - \frac{\cos k_v}{k_v} \right) \right)$$
(8.21)

where R_1 is the truncation error. We shall consider the typical magnitude of R_1 in the next section, but for the moment we set $R_1 = 0$. With this approximation in place, we then have a model with two observations (the amplitude and phase of $\tilde{\gamma}$) and three unknowns: $\phi_0, k_v = \beta_z h_v/2$, which is a function of the unknown height h_v and known baseline β_z , and a_{10} , a normalized linear Legendre coefficient. Hence we have more unknowns than observations, and so in order to be able to invert the model we require additional information. Varying the polarisation, while keeping the wavelength, sensor geometry and baseline constant, provides a convenient way to add measurement diversity without adding new parameters. Indeed, it is reasonable to assume that β_v , ϕ_0 , f_0 and f_1 all remain invariant to changes in polarisation and only the Legendre coefficient a_{10} can change, reflecting changes in the structure function with polarisation. In the most general case we can then consider adding several polarisation channels to the model of equation (8.21), but in the simplest we require just two, with scattering mechanisms \underline{w}_1 and \underline{w}_2 , providing four observations and four unknowns, as shown in the following pair of equations:

$$\tilde{\gamma}\left(\underline{w}_{1}\right) = e^{i(k_{v}+\phi_{o})}(f_{0}+a_{10}\left(\underline{w}_{1}\right)f_{1}) = e^{i\phi}(f_{0}+a_{10}\left(\underline{w}_{1}\right)f_{1})$$

$$\tilde{\gamma}\left(\underline{w}_{2}\right) = e^{i(k_{v}+\phi_{o})}(f_{0}+a_{10}\left(\underline{w}_{2}\right)f_{1}) = e^{i\phi}(f_{0}+a_{10}\left(\underline{w}_{2}\right)f_{1})$$
(8.22)

This is a more balanced set suitable for inversion; that is, for estimation of the four unknown parameters ϕ_0 , k_v , $a_{10}(\underline{w}_1)$ and $a_{10}(\underline{w}_2)$ from two observations of complex coherence. The following strategy then follows immediately from equation (8.22). First we estimate the surface phase term (noting that f_1 is imaginary), not by line fitting as in equation (8.5), but by differencing the complex coherences as shown in equation (8.23):

$$\phi = k_v + \phi_o = \arg(-i(\tilde{\gamma}(\underline{w}_1) - \tilde{\gamma}(\underline{w}_2))) \tag{8.23}$$

We can then calculate k_v from the real part of the phase-shifted coherence, as shown in equation (8.24):

$$\operatorname{Re}(\tilde{\gamma}(\underline{w}_1)e^{-i\phi}) = \operatorname{Re}(\tilde{\gamma}(\underline{w}_2)e^{-i\phi}) = \frac{\sin k_v}{k_v} \quad 0 \le k_v \le \pi$$
(8.24)

Note that to invert this relation to estimate k_v we can use the following convenient invertible approximation for the SINC function, valid over the range 0 to π :

$$y = \frac{\sin(x)}{x} \approx \sin\left(\frac{(\pi - x)}{2}\right)^{1.25} \quad 0 \le x \le \pi, \ 0 \le y \le 1$$
$$\Rightarrow x \approx \pi - 2\sin^{-1}(y^{0.8}) \quad (8.25)$$
$$\Rightarrow \hat{k}_y \approx \pi - 2\sin^{-1}(\operatorname{Re}(\tilde{\gamma}(\underline{w}_2)e^{-i\phi})^{0.8})$$

Note that this approach, when combined with equation (8.23), allows us to calculate the surface phase ϕ_0 without the need for a separate straight-line coherence region assumption. We can then calculate the structure parameter for arbitrary polarization \underline{w} from the imaginary part of the phase-shifted coherence, as shown in equation (8.26):

$$\hat{a}_{10}(\underline{w}) = \frac{\text{Im}(\tilde{\gamma}(\underline{w}) e^{-i\phi})}{|f_1|} = \frac{\text{Im}(\tilde{\gamma} e^{-i(k_v + \phi_0)})k_v^2}{\sin k_v - k_v \cos k_v}$$
(8.26)

Finally, we can reconstruct the vertical profile by knowing the interferometric wavenumber β_z to calculate the height from k_v and using the Legendre coefficient to reconstruct the profile with unit integral over this height range, as shown in equation (8.27):

$$\hat{h}_{\nu} = \frac{2\hat{k}_{\nu}}{\beta_{z}} \Rightarrow \hat{f}_{L1}(z) = \frac{1}{\hat{h}_{\nu}} \left((1 - \hat{a}_{10}) + \frac{2\hat{a}_{10}}{\hat{h}_{\nu}} z \right) \quad 0 \le z \le \hat{h}_{\nu}$$
(8.27)

Figure 8.7 shows a schematic summary of the types of structure function we can construct from this simple first-order truncation. We note that the maximum and minimum of this first-order structure function are given simply in terms of the parameter a_{10} , as shown in Figure 8.7. Note that for $a_{10} > 1$ this has a negative minimum on the surface. This may seem to violate the important physical requirement that f(z) be non-negative (since physically it represents scattered power as a function of depth). However, such a restriction is not necessary when



Fig. 8.7 Summary of first-order reconstruction of the vertical structure function



Fig. 8.8 Examples of bipolar Legendre structure function estimates of a non-negative step structure function

we realize that $f_{\rm L}(z)$ is only a band-limited approximation to the true structure function. While the true function is always non-negative, its approximation can go negative, indicating more concentrated scattering from the top of the layer. To illustrate this we show, in Figure 8.8, various Legendre approximations (up to sixth order) for a vertical step function at 50% of layer depth; that is, the true structure function involves uniform scattering, but only from the top half of the layer. We see that all approximations (including the first-order as used in $f_{\rm L}(z)$ go negative at low z values, and this is a direct consequence of the true physical structure. Hence it is useful to allow such negative profiles in the estimation, on the understanding that we only ever obtain an approximation to the true profile. If it is important to maintain positivity in the approximation, then the negative parts of the profile can be set to zero and the estimate still have some physical correspondence with the true profile. In this case, for the linear approximation of this model, the scattering is non-zero only for elevated heights above (below) a critical height, z_c , expressed simply in terms of a_{10} , as shown in equation (8.28):

$$|a_{10}| > 1 \Rightarrow z_c = \frac{(a_{10} - 1)}{2a_{10}} h_v = \frac{h_v}{2} \left(1 - \frac{1}{a_{10}} \right)$$
(8.28)

For positive a_{10} , z_c lies between the surface and half the height, while for negative it lies between the top and the half height.

Before proceeding further it is important to consider the range of the three unknown parameters ϕ_0 , k_v and a_{10} with a view to investigating the uniqueness of this model inversion. The phase ϕ_0 is defined in the range 0 to 2π , with ambiguities arising for surface variations in excess of this. These correspond to the classical phase unwrapping problem in radar interferometry. However, here we are more concerned with phase shifts relative to ϕ_0 , and can therefore ignore the phase unwrapping problem—at least initially. If we consider scenarios where β_z is always positive (by appropriate selection of master and slave tracks for the baseline generation) and h_v is also positive, then a good working range for k_v is $0 \le k_v \le \pi$. Although, mathematically, k_v can go to infinity, in practice we would wish to restrict it to avoid phase ambiguities in the layer; that is, we design the interferometer to ensure that the depth of the layer is always less than the 2π ambiguity height of the interferometer. This then restricts k_v to the range specified.

The Legendre coefficient a_{10} , on the other hand, can be bipolar—negative or positive—but must be constrained so that the magnitude of the right-hand side in equation (8.21) is always less than or equal to 1 (to match the limits of coherence on the left-hand side). This requires the following inequality to hold:

$$|\tilde{\gamma}| \le 1 \Rightarrow |a_{10}| \le \sqrt{\frac{1 - f_0^2}{|f_1|^2}} = \sqrt{\frac{k_v^2 - \sin^2 k_v}{\frac{\sin^2 k_v}{k_v^2} - 2\frac{\sin k_v \cos k_v}{k_v} + \cos^2 k_v}}$$
(8.29)

The range of a_{10} allowed under this constraint varies as a function of k_v , as shown in Figure 8.9. Note that the range is limited to ± 1.732 for low values, and up to $\pm \pi$ for large values of k_v . Note also that the variation of coherence with structure (k_v fixed and a_{10} variable) is a straight line in the complex plane, intersecting the unit circle at two points corresponding to the limits given in equation (8.29).

It is interesting to look at the variation of this line with k_v . We show such a loci in Figure 8.10, where we have removed any topographic phase so that the surface phase is zero and lies at the point O. The set of lines are for variation of k_v from 0° to 180° in 5-degree steps. The loci of points of the uniform SINC model ($a_{10} = 0$) are shown as black stars, which we see constitute a spiral. However, passing through each SINC value is now a straight line (solid for positive a_{10} values, and dashed for negative). Clearly, if ϕ_0 is not zero then the whole diagram is just rotated clockwise for negative and anticlockwise for



Fig. 8.9 Maximum (dash) and minimum (solid) bounds on first Legendre coefficient as a function of k_v



Fig. 8.10 Coherence loci for first-order Legendre approximation: positive a_{10} (solid) and negative a_{10} (dash) for different k_v values (black stars)

positive phase shifts. In any case there are two important observations from this result:

- 1. The first-order model does not cover the whole coherence diagram, and severely limits the possible set of valid coherences. In fact, for a fixed k_v the valid coherences are constrained to lie along a single straight line going through the appropriate SINC point.
- 2. It follows from this model that the coherence variation with polarisation should also lie along a line in the complex plane. However, this line does not intersect the unit circle at the topographic phase point, and hence is not the same line as used in other coherence models, such as the two-layer random-volume-over-ground or RVOG model.

We also note from Figure 8.10 that there is also some ambiguity for phase shifts less than $\pi/2$, where we see the intersection of different lines. However, these ambiguities can be explained physically as the equivalence of a thick layer with all scattering from the surface region having the same phase centre as a thin layer with all scattering coming from the top. In order to enable a unique inversion we can restrict the model to positive values of a_{10} ; that is, to scenarios where the scattering increases rather than decreases with height into layer 1. These positive loci are shown in Figure 8.11, where we also superimpose the linear coherence loci of RVOG and its distortion for temporal effects on volumeonly scattering. Clearly, from this the first-order Legendre series cannot be used to fully represent coherences with large μ values in the RVOG model. Nor can it even be used for all temporal decorrelations in the volume-only scattering case.

From these observations we conclude that the conditions are very restrictive for the inversion scheme of equations (8.23)–(8.27) to apply. Consequently, we can consider the first-order Legendre model to be inappropriate for general physical applications. At the very least, for fixed k_v , we require better coverage



Fig. 8.11 Overlay of RVOG model on positive first-order Legendre model

of the complex plane so that we can model a wider range of physical scenarios (including the RVOG model). To do this we must extend the model to second order, as we now consider. However, we shall see that in doing so we can still make use of some of the inversion ideas from this first-order truncation.

8.2.2 Second-order Legendre model

We now consider truncation of the Legendre series to second order, as shown in equation (8.30). Again we assume that we can isolate all polarisation dependence in the Legendre spectrum, which amounts to the assumption that height, topographic phase and baseline are all invariant to polarisation changes.

$$\tilde{\gamma}(\underline{w}) = e^{i(k_v + \phi_o)} \left\{ \frac{\sin k_v}{k_v} + a_{10}\left(\underline{w}\right) i\left(\frac{\sin k_v}{k_v^2} - \frac{\cos k_v}{k_v}\right) + a_{20}(\underline{w}) \left(\frac{3\cos k_v}{k_v^2} - \left(\frac{6 - 3k_v^2}{2k_v^3} + \frac{1}{2k_v}\right)\sin k_v\right) \right\} + R_2$$

$$(8.30)$$

Here R_2 is again the truncation error. This error is of the order of the absolute value of the next term in the coherence series normalized by $|f_0|$; in this case, $R_2 \approx |f_3|/|f_0|$. This is generally small. From Figure 5.15 we see that for $0 \le k_v \le \pi$ this is much smaller than the first-order truncation error R_1 . Hence the second-order truncation offers a much more accurate model; but it has an increased number of parameters, and so we now turn to consider its suitability for inversion.

We see that we now have two polarisation dependent coefficients, a_{10} and a_{20} , which together with k_v and ϕ_0 constitute a set of four unknowns to be determined from the two observations (amplitude and phase of coherence). Here, polarisation diversity does not seem to help us, as each additional <u>w</u> adds two observations but also adds two new unknowns. Instead we need to develop a different strategy to invert equation (8.30). First, however, we investigate the expanded coverage in the complex plane of this second-order model.

To visualize coverage of the second-order Legendre coherence model (equation (8.30)) inside the unit circle, we rewrite the coherence in the following second-order form:

$$\begin{split} \tilde{\gamma}(\underline{w}) &= e^{i(k_{v} + \phi_{o})}(f_{0} + a_{10}(\underline{w})f_{1} + a_{20}(\underline{w})f_{2}) \\ &= z(a_{20}) + a_{10}d \\ &\Rightarrow d = e^{ik_{v} + \phi}i\left(\frac{\sin k_{v}}{k_{v}^{2}} - \frac{\cos k_{v}}{k_{v}}\right) \end{split}$$
(8.31)
$$z(a_{20}) &= e^{ik_{v} + \phi}\left(\frac{\sin k_{v}}{k_{v}} + a_{20}\left\{\frac{3\cos k_{v}}{k_{v}^{2}} - \left(\frac{6 - 3k_{v}^{2}}{2k_{v}^{3}} + \frac{1}{2k_{v}}\right)\sin k_{v}\right\}\right) \end{split}$$

In this form we see that a_{10} again generates a line in the complex plane passing through a fixed-point z with direction d, but that now the fixed point is itself determined by the second-order coefficient a_{20} . Since the function f_2 is always negative (see Figure 5.15), it follows from equation (8.31) that for *negative* a_{20} the fixed point moves radially *outwards* towards the unit circle, and for *positive* a_{20} it moves *inwards* towards the origin of the coherence diagram. Thus, for a fixed k_v the result is a family of straight lines, all with the same slope and defined by a set of fixed points generated by a radial line passing through the SINC point, as shown in Figure 8.12. Here we show in thick black line the spiralling SINC locus for $\phi_0 = 0$. For each point on this curve there is now a family of lines generated for positive and negative a_{10} (shown as solid and dashed lines respectively) and moving up and down the radial line through the SINC point according to positive or negative a_{20} . As we move along the SINC locus this pattern of lines is rotated and shifted accordingly. The bounds of a_{10} for fixed a_{20} can then be derived by setting the coherence to unity, as shown in



Fig. 8.12 Coherence loci for the secondorder Legendre approximation equation (8.32):

$$\begin{split} |\tilde{\gamma}|^2 &= (f_0 + a_{10}f_1 + a_{20}f_2) (f_0 + a_{10}f_1 + a_{20}f_2)^* \\ &= f_0^2 + 2a_{20}f_0f_2 + a_{10}^2f_1^2 + a_{20}^2f_2^2 = 1 \\ &\Rightarrow -\sqrt{\frac{1 - (f_0 + a_{20}f_2)^2}{|f_1|^2}} \le a_{10} \le \sqrt{\frac{1 - (f_0 + a_{20}f_2)^2}{|f_1|^2}} \end{split}$$
(8.32)

If we then set $a_{10} = 0$ we obtain the corresponding bounds of a_{20} , as follows:

$$(f_0 + a_{20}f_2)^2 = 1 \Rightarrow \frac{1 - f_0}{f_2} \le a_{20} \le -\frac{(1 + f_0)}{f_2}$$
 (8.33)

We note that for a fixed k_v value these bounds now lead to *full* coverage of the unit circle. Hence, if we know the k_v value then we can use the position of *any* sample coherence to estimate the two parameters a_{10} and a_{20} , as shown in equation (8.34):

$$a_{10}(\underline{w}) = \frac{\operatorname{Im}(\tilde{\gamma}(\underline{w}) e^{-i\phi})}{|f_1|} \qquad a_{20}(\underline{w}) = \frac{\operatorname{Re}(\tilde{\gamma}(\underline{w}) e^{-i\phi}) - f_0}{f_2} \qquad (8.34)$$

The structure function itself can then be expressed as in equation (8.35):

$$\hat{f}_{L2}(\underline{w},z) = \frac{1}{h_{\nu}} \left(1 - \hat{a}_{10} + \hat{a}_{20}(\underline{w}) + \frac{2z}{h_{\nu}} (\hat{a}_{10}(\underline{w}) - 3\hat{a}_{20}(\underline{w})) + \hat{a}_{20}(\underline{w}) \frac{6z^2}{h_{\nu}^2} \right) \quad 0 \le z \le h_{\nu}$$
(8.35)

One interesting property of the second-order Legendre approximation $\hat{f}_{2L}(z)$ is that its extreme value (maximum or minimum) no longer has to fall at the boundaries of the layer. This provides us with more flexibility in representing variations in the structure function itself, which is important in complex media such as scattering from forest canopies (Woodhouse, 2006). The stationary point of the estimated profile can be simply related to the Legendre coefficients as shown in equation (8.36):

$$\frac{df_{L2}(z)}{dz} = 0 \Rightarrow z_m = h_v \left(\frac{1}{2} - \frac{a_{10}}{6a_{20}}\right)$$
(8.36)

Here we see, for example, if $a_{10} = 0$ then the minimum (maximum) of the profile occurs at half the layer depth for positive (negative) a_{20} . In general, if the ratio a_{10}/a_{20} is positive then the extreme point will occur in the lower half of the layer, while for a negative ratio the extreme point will occur in the upper half. In this way we can represent a much wider variety of structure functions than is possible using the classical RVOG model (which assumes that the maximum always occurs at the top of the volume). In particular we note that since a_{20} can be positive or negative, we can now represent functions with a maximum response below the top of the layer (negative a_{20}) or with an enhanced response from the surface position at z = 0 (a_{20} positive). The former is useful for representing non-exponential volume scattering profiles, while the latter can be used to represent changes in μ —the effective surface-to-volume scattering ratio.

To see an example of the flexibility of this second-order approximation we consider its application to the two-layer RVOG model (see Section 7.4.2). The special form of this model that we use is summarized in equation (8.37):

$$\tilde{\gamma}\left(\underline{w}\right) = e^{i\phi_0} \frac{\tilde{\gamma}_v + \mu\left(\underline{w}\right)}{1 + \mu\left(\underline{w}\right)} \quad \tilde{\gamma}_v = e^{ik_v} \frac{\sin k_v}{k_v} \tag{8.37}$$

For simplicity we show a case where the volume-only coherence ($\mu = 0$) is given by a simple zero extinction medium (a uniform structure function). The factor μ then corresponds physically to the ratio of surface-to-volume scattering. We also assume that the surface phase $\phi_0 = 0$ in this example. Figure 8.13 shows how this model maps onto the Legendre coordinates (defined by k_v) for μ in the range -30 dB to +30 dB in 1-dB steps.

Here we have used a specific example when $h_v = 10$ m, $\beta_z = 0.2$, and so $k_v = 1$. Each point of the RVOG model now has a set of Legendre coordinates a_{10}, a_{20} , as shown in Figure 8.14. We see that when μ is small the coordinates are both zero, corresponding to the assumption of a uniform structure. However, as μ increases we see that a_{10} increases in a negative direction while a_{20} increases



Fig. 8.13 The RVOG model (stars) superimposed on the second-order Legendre coordinate system



Fig. 8.14 Variation of Legendre coefficients a_{10} (solid) and a_{20} (dash), with μ for the RVOG model



Fig. 8.15 Variation of second-order Legendre structure function approximation of the RVOG model with surface-to-volume scattering ratio μ

in the positive direction. This reflects changes in the structure function itself. Figure 8.15 shows an image of how the second-order approximation to the RVOG structure function varies with μ . Each vertical profile extends over 10 m, and is normalized so that its integral is unity (as in equation (8.35)). On the left we see the uniform volume scattering profile obtained when $\mu = 0$. As μ increases we see a shift in the structure function to more localized surface scattering, as physically expected in the RVOG model. Again we note that as the surface contribution increases, the second-order approximation is forced to go negative at some points in the volume. This again reflects the approximate nature of the truncation rather than any physical interpretation of negative scattering amplitudes.

In conclusion, we have seen that a second-order Legendre expansion is the lowest-order truncation capable of providing full unit circle coverage. We have demonstrated its application to the widely used RVOG coherence model to demonstrate its ability to reflect changes in the underlying structure function. There remains, however, one issue with this model. It has too many unknowns to be inverted and hence to be directly applied for height estimation. In the next section we turn to consider methods for resolving this limitation.

8.2.3 Approximate height estimation from the second-order Legendre series

In the previous section we showed that a second-order truncation of the Legendre series is useful for characterising a wide range of different structure functions. The problem with this model is that we have more unknowns than observations. For a single polarization channel, single wavelength and single baseline we have only two observations (one complex coherence), while the model has four unknowns (the two Legendre coefficients and two structural parameters ϕ_0 and k_v). The only observation in our favour is that ϕ_0 and k_v are invariant to changes in polarisation. In order to progress, therefore, we need somehow to estimate two of these parameters so as to obtain a balance of two unknowns and two observations. In this section we consider approximate methods for achieving this. In particular we look at ways of estimating ϕ_0 and k_v , with layer depth then following from the latter. To do this we will need to impose some further constraints on properties of the unknown structure function, but we will see that these can still be rather lax, allowing some flexibility (and critically more so than in the fixed structure approaches like RVOG) in determining variations in structure.

We start by noting that if we adopt the slightly more general SVOG model for our second-order Legendre series (see Section 7.4.3) we can still use a line-fit technique to obtain estimates of ϕ_0 ; that is, the surface phase can again be estimated from equation (8.5) or (8.16). In this scheme we maintain the assumption that the upper layer comprises random volume scatterers; the only difference with RVOG is that the volume contribution can now have arbitrary structure function and not just an exponential.

In order to estimate height we first use this ϕ_0 estimate to obtain a phase-based estimate, exactly as proposed in equation (8.19). However, as we noted earlier, this phase centre separation, according to SVOG, can lie anywhere between halfway and the top of the layer, and hence in general underestimates the true layer depth.

To progress, one key idea is that this error can be at least partly compensated by employing a coherence amplitude correction term. The idea is that as the phase centre separation increases due to changes in structure function so, *at the same time*, the effective volume depth decreases (as the structure function becomes more localized near the top of the layer), and hence the level of volume decorrelation will decrease. A convenient invertible model for this coherence amplitude process is just the f_0 or SINC coherence function, as discussed in equation (8.25). Just as required, when coherence amplitude decreases, so this height (or k_v) estimate will decrease at the same time as the phase estimate increases. Finally, by combining these two terms with a scaling parameter η we then obtain an approximate algorithm that can compensate variations in structure, as shown in equation (8.38):

$$\hat{k}_{\nu} = \frac{1}{2} \left\{ \arg(\tilde{\gamma}_{w_{\nu}} e^{-i\hat{\phi}_{0}}) + \eta(\pi - 2\sin^{-1}(|\tilde{\gamma}_{w_{\nu}}|^{0.8})) \right\}$$
$$\Rightarrow \hat{h}_{\nu} = \frac{2\hat{k}_{\nu}}{\beta_{\tau}}$$
(8.38)

The first term represents the phase component (using the estimated surface phase together with our estimate of 'volume-only' coherence channel). The second—the coherence amplitude correction—is weighted by η , to be selected so as to make the full expression as robust as possible to changes in the structure function.

This expression has the right kind of behaviour in two important special cases. If the medium has a uniform structure function then the first term will give half the height or $\beta_z h_v/2$, but the second will then also obtain half the true height and yield $\beta_z h_v/2$ (if we set $\eta = 1$), and so half the sum gives the correct k_v estimate. At the other extreme, if the structure function in the volume channel is localized near the top of the layer, then the phase height will give

the true height $\beta_z h_v$, and the second term will approach zero. Half the sum then still produces the correct k_v estimate. The idea is that equation (8.38) will provide a reasonable estimate for arbitrary structure functions between these two extremes. It requires estimates of only two parameters: the true surface topographic phase ϕ_0 , and the volume-only complex coherence $\tilde{\gamma}(w_v)$.

We can extend this idea further and estimate an optimum value of weighting factor η by using the second-order Legendre structure model for the volume coherence channel, as shown in equation (8.39):

$$\tilde{\gamma}(\underline{w}_{\nu}) = e^{i(k_{\nu} + \phi_o)} (f_0 + a_{10} \left(\underline{w}_{\nu}\right) f_1 + a_{20} \left(\underline{w}_{\nu}\right) f_2)$$
(8.39)

Now, if we allow a_{10} and a_{20} their full range we can fit this coherence to any k_v value, and hence seem to undermine the approximation proposed in equation (8.38). However, by making some reasonable physical assumptions about volume scattering we can reduce the working range of a_{10} and a_{20} as follows.

The basic idea is that in the selected special volume-only channel we assume there is zero surface scattering component, and hence its structure function should have a local minimum at the surface (at z = 0). This in turn requires that in this polarization channel \underline{w}_v we restrict $a_{10} \ge 0$ and $a_{20} \le 0$. When combined with the limits derived in equations (8.32) and (8.33) we will see that this constrains the k_v values satisfying equation (8.39). Figure 8.16 illustrates the results. Here we show along the abscissa a set of true k_v values in steps of 0.1 over the range 0–2. The corresponding ordinate shows the spread of estimated k_v values obtained using equation (8.38) with $\eta = 0.8$. This value is selected to fit the $a_{20} = 0$ variation, and the underestimates we see are then entirely due to the presence of non-zero a_{20} structure. It also ensures that equation (8.38) will always estimate the minimum height consistent with the data. While the trend in Figure 8.16 is encouraging, some of the errors can apparently be quite large. For example, if the estimate yields a value of 1 (in the ordinate of Figure 8.16)



Fig. 8.16 Estimated versus true k_v values for the full range of volume structure functions a_{10}, a_{20} , and using $\eta = 0.8$ in equation (8.39)



Fig. 8.17 Fractional error for full range of volume structure functions for $k_v = 2$

then we see that the true value could actually lie anywhere between 1 and 1.6, depending on the volume structure. However, this simple interpretation masks an important issue. The underestimation errors for each abscissa point in Figure 8.16 increase in proportion to a_{20} , as shown for the $k_v = 2$ case in Figure 8.17. Here we plot the fractional error in k_v estimate for all valid values of a_{10} and a_{20} (but note that this behaviour is typical of all other values, and so our conclusions will apply equally for all values of k_v).

The key observation from Figure 8.17 is that the largest errors always occur for large a_{20} values. This makes physical sense, as this occurs for a quadratic profile, which is a second-order approximation to a Dirac delta function located halfway up the volume. In this case we can achieve a combined high-coherence and low-phase centre, in contradiction to the assumptions behind equation (8.38). This is a problem faced by all height estimation techniques based on interferometry. If the volume has a top height h_v but the bulk of the scattering comes from halfway up the volume, then interferometry 'sees' a smaller effective height. In order to resolve such ambiguities we need to add extra information beyond a single baseline single wavelength interferometer. Other possibilities include adding more baselines or using a second frequency (Treuhaft, 2000b, 2004; Reigber, 2000, 2001; Neumann, 2008).

If we accept that such errors can occur for a single baseline, but only for rather extreme (and unlikely) cases of the structure function, then we can proceed to employ equation (3.38) as a reasonable approximation, with around 10–15% estimation errors for a wide range of structure function variations.

We now turn to consider depth retrieval under the more extreme assumption of a fixed structure function: namely, the exponential assumed by the RVOG class of models.

8.2.4 Height estimation using RVOG

The above algorithm assumes that the volume-only structure function $f_v(z)$ has the property that its Fourier–Legendre coherence contributions satisfy the

bounds $a_{10} \ge 0$, $a_{20} \le 0$. To avoid such assumptions about the Fourier– Legendre expansion we can instead construct a similar approach, but based directly on the two-parameter RVOG model, where mean wave extinction replaces the Legendre structure parameter a_{10} . This automatically generates a structure function that decreases with depth. The layer depth is now obtained by minimising the following function $G(\lambda)$, being the norm of the difference between the volume coherence (itself obtained from the observed coherence $\tilde{\gamma}_{\underline{w}_{\nu}}$ by a line shift through the parameter $0 \le \lambda \le 1$) and the RVOG model prediction for volume-only scattering ($\mu = 0$). The phase ϕ_2 is the second unit circle intersection point of the straight line fit (see Figure 8.2).

$$\min_{h_{\nu},\sigma} G\left(\lambda\right) = \left\| \tilde{\gamma}_{\underline{w}_{\nu}} + \lambda \left(e^{i\hat{\phi}_{2}} - \tilde{\gamma}_{\underline{w}_{\nu}} \right) - e^{i\hat{\phi}} \frac{p}{p_{1}} \frac{e^{p_{1}h_{\nu}} - 1}{e^{ph_{\nu}} - 1} \right\| \quad \text{where} \begin{cases} p = \frac{2\sigma_{e}}{\cos\theta} \\ p_{1} = p + i\beta_{z} \\ \beta_{z} = \frac{4\pi\Delta\theta}{\lambda\sin\theta} \end{cases}$$
$$\min_{h_{\nu},\sigma} G\left(\lambda = 0\right) = \left\| \tilde{\gamma}_{\underline{w}_{\nu}} e^{-i\hat{\phi}} - \frac{p}{p_{1}} \frac{e^{p_{1}h_{\nu}} - 1}{e^{ph_{\nu}} - 1} \right\| \tag{8.40}$$

Shown in the lower portion of equation (8.40) is the simplified version of this algorithm obtained in the case $\lambda = 0$; that is, when we can assume that the observed coherence $\tilde{\gamma}_{\underline{W}_{\nu}}$ itself has $\mu = 0$ —that it contains volume-only scattering. We have also shifted the topographic phase estimate onto the observable coherence. Again it is interesting to investigate the range of the two parameters h_{v} and σ_{e} with a view to determining coverage and uniqueness of the solution in the complex plane. The range of h_{v} is again capped by the requirement to avoid phase ambiguities in the layer, so that $0 \le \beta_{z} h_{\nu} \le 2\pi$. The extinction can also have non-negative infinite extent; but in practice, as extinction becomes large enough so the coherence amplitude becomes insensitive to changes in height, and varies only the phase of the top of the layer. These trends are apparent in the loci shown in Figure 8.18. Here we see the variation of coherence over



Fig. 8.18 Overlay of SVOG linear region on RVOG coherence loci

the full range of height for varying extinction. The inside spiral corresponds to the reference zero extinction or SINC loci. We see that for zero extinction the coherence falls to zero at the 2π height. However, as the extinction increases, the curvature of the spiral reduces until for large extinction the loci is almost circular, maintaining high coherence amplitude because of the small effective volume contributing to decorrelation.

Although the loci are no longer straight lines we note that they provide a set of non-intersecting curves, and so again if we overlay a sample volume coherence (shown as the point in Figure 8.18) then we can find a unique solution to equation (8.40) (for fixed λ), and thus secure an estimate of h_v . Note again that here the extinction parameter is acting as a structure compensation parameter, allowing height estimation for a wide range of structure functions approximated by exponentials with varying extinction rates. Note also that the coverage is not complete. If we draw a line through our sample coherence (shown as the point in Figure 8.18) we see that if μ is above a certain level the coherence can still fall outside coverage of the simple extinction model. In this case we have to employ the free parameter λ in equation (8.40) to move the coherence back into the valid region, but generally we have no idea which value of λ to use and hence are likely to make errors in the height retrieval. In this sense we see that the SINC curve defines a boundary for coverage, and any volume scattering candidate coherence must lie above the SINC curve to enable a clear solution in RVOG inversion.

The above algorithm uses the RVOG model to match the observed coherence in both amplitude and phase. However, to do this we require estimates of the topographic phase ϕ_0 . We saw in Section 8.1 how to devise algorithms based on RVOG and OVOG to estimate this parameter. However, sometimes the coherence can be so low that this phase estimate is too noisy to use. In this case we would like to employ an algorithm that does not make use of phase information and relies only on coherence amplitude. We can devise such a model based on the RVOG approach, as long as we assume that the structure function has a known form (known extinction, in this case). We then have only one unknown (the depth of the layer) and one observable, and can solve a minimization problem as shown in equation (8.41):

$$\min_{h_{\nu}} G = \left\| \left| \tilde{\gamma}_{\underline{w}_{\nu}} \right| - \left| \frac{p}{p_1} \frac{e^{p_1 h_{\nu}} - 1}{e^{p h_{\nu}} - 1} \right| \right\| \text{ where } \begin{cases} p = \frac{2\sigma_e}{\cos \theta} \\ p_1 = p + i\beta_z \end{cases}$$
(8.41)

In the RVOG case, a known structure function implies knowledge of the mean extinction coefficient in the medium $\overline{\sigma}_e$, as shown in equation (8.41). This can sometimes be estimated from physical models of the environment or from measurements and inversions from previous datasets (see Section 3.5.3). Note, however, that matching coherence amplitude calls for good calibration and compensation for SNR and temporal effects.

In conclusion, we have seen three important algorithms for estimating layer depth (or height) from single baseline polarimetric interferometric data. Equation (8.38) represents an approximate method that makes minimal assumptions about the layer structure function. Equation (8.40) assumes an exponential structure via the RVOG model, and consequent matching of the coherence in both amplitude and phase provides an estimate of both height and mean

extinction. Finally, if surface phase estimates are not available, a coherence amplitude-only approach is shown in equation (8.41). This, however, requires an *a priori* estimate of the mean extinction in the medium.

8.2.5 Depth estimation using OVOG

We turn now to consider problems faced in estimating layer depth h_v when layer 1 behaves as an oriented volume, with polarisation-dependent propagation through the layer. In this case we can no longer consider having just a single volume coherence point for inversion but a triplet of such points, as discussed in Section 7.4.5. Nonetheless, we can still use the techniques developed for the SVOG and RVOG models to derive simple algorithms for depth estimation in oriented volumes, as follows (Treuhaft, 1999; Cloude, 2000a; Lopez-Sanchez, 2006, 2007).

In the same way as for topography estimation, we consider two special forms of the oriented volume-over-surface model: the first with assumed small surface contributions in all channels (to be called the finite OV problem), and the secondly with high μ dispersion, when only one channel (with the highest extinction) has a volume-dominated response while the other (the lowest extinction) has a large surface-to-volume ratio μ . We now consider each of these cases in turn.

8.2.5.1 Finite OV height estimation algorithm

In this case we deal with backscattering by a cloud of volume scatterers, such that there is no surface scattering from top or bottom of the layer, as shown schematically in Figure 8.19. We are then interested in an algorithm for deriving the layer depth and bottom phase from the triplet of observed coherences. As the medium is not infinite we cannot use the OV solution of equation (8.17) to find bottom or top phase by a circle fit. We therefore need to consider a more integrated parameter estimation problem, whereby bottom phase is included as an unknown at the same time as layer depth.

8.2.5.2 First-order Legendre OV inversion

We start by noting that despite the complications of anisotropic propagation, all three coherences are characterized by the same depth h_v (and hence the same k_v value) and surface phase ϕ , and differ only in their structure functions. So, by assuming $\mu = 0$ for all channels we obtain the ordered triplet of volumeonly coherences $\tilde{\gamma}_{xx} = \tilde{\gamma}_1, \tilde{\gamma}_{xy} = \tilde{\gamma}_2, \tilde{\gamma}_{yy} = \tilde{\gamma}_3$ from the eigenpolarisations x and y, to set up the following Legendre cost function, based on that derived in



Fig. 8.19 Geometry of scattering from an oriented volume with finite depth


Fig. 8.20 Example oriented volume coherence triplet superimposed on coherence loci for a first-order Legendre structure function

equation (8.21).

$$\min_{k_{\nu},a_{1},a_{2},a_{3}}G = \sum_{j=1}^{3} \left\| \tilde{\gamma}_{j} - e^{i\phi}e^{ik_{\nu}} \left(\frac{\sin k_{\nu}}{k_{\nu}} + ia^{j} \left(\frac{\sin k_{\nu}}{k_{\nu}^{2}} - \frac{\cos k_{\nu}}{k_{\nu}} \right) \right) \right\|$$
(8.42)

Note that here there are five unknowns and six observables. The parameter bounds are the same as those derived earlier $(0 \le k_v \le \pi, 0 \le a^j \le \pi, 0 \le \phi < 2\pi)$, and the only difference now is that for each fixed pair of values k_v, ϕ we seek a triplet of Legendre coefficients a^j that minimize the above function. The global minimum of such searches will then give us estimates of the parameters of the layer. There is, however, a simple geometrical interpretation of this inversion, as we now consider.

The first point to make is that the model of equation (8.42) implies that the triplet of coherences lie along a line in the complex plane (see Figure 8.20). These lines are shown in Figure 8.11, and derive from the assumption of a truncated Legendre series. Hence a starting point for the suitability of the inversion of (8.42) is to test whether or not the components of the triplet are collinear. If they are, then equation (8.42) will have a good match with the data, otherwise the assumptions of the truncated Legendre series may be invalid. Note, incidentally, that this line does not intersect the unit circle at the surface topography point. This is in contrast to the RVOG line fit employed in Section 8.1 to find surface topography from this intersection. It is therefore important in applications to be able to differentiate between RVOG and finite OV before applying the appropriate parameter estimation. This can be accomplished in several ways—for example, by checking the orthogonality of the optimum coherence states: for single layer OV they will be orthogonal, while for two-layer RVOG they will not be orthogonal.

The second point to note is that the slope of the line joining the three coherences is given simply in terms of the two fixed parameters k_v and ϕ , as shown in equation (8.43). Therefore, by fitting a line through the three volume coherences

and measuring its slope, we can obtain directly an estimate of the parameter $k_v + \phi$. We can then employ a single channel model fit—as in the random volume case—to any one of the three coherences (for example, the maximum coherence γ_1) to obtain an estimate of k_v from minimization of the function shown in equation (8.43).

$$\begin{split} \tilde{\gamma} &= e^{i\phi} e^{ik_v} \left(\frac{\sin k_v}{k_v} + ia^j \left(\frac{\sin k_v}{k_v^2} - \frac{\cos k_v}{k_v} \right) \right) = z + a^j d \\ \Rightarrow d &= e^{ik_v + \phi} i \left(\frac{\sin k_v}{k_v^2} - \frac{\cos k_v}{k_v} \right) \\ \Rightarrow m &= -\frac{1}{\tan(k_v + \phi)} \end{split}$$
(8.43)
$$\Rightarrow (k_v + \phi) &= -\tan^{-1} \left(\frac{1}{m} \right) \\ \Rightarrow \min_{k_v, a^1} G &= \left\| \tilde{\gamma}_1 - e^{i\phi + k_v} \left(\frac{\sin k_v}{k_v} + ia^1 \left(\frac{\sin k_v}{k_v^2} - \frac{\cos k_v}{k_v} \right) \right) \right\| \end{split}$$

This value of k_z will then, by definition, satisfy the other two polarisation channels, and when combined with the slope estimate will provide an estimate for the surface topography ϕ .

Having determined a simple algorithm based on the Legendre coherence expansion, we now consider a solution based on exponential structure functions.

8.2.5.3 Exponential OV inversion

The differences in structure function with polarisation may be ascribed to exponentials, as in the OVOG model. In this case we can derive a new cost function as shown in equation (8.44):

$$\min_{h_{\nu},\kappa_{1},\kappa_{2}} G = \sum_{j=1}^{3} \left\| \tilde{\gamma}_{j} - e^{i\phi} \frac{p_{j}}{(p_{j} + i\beta_{z})} \frac{e^{(p_{j} + i\beta_{z})h_{\nu}} - 1}{e^{p_{j}h_{\nu}} - 1} \right\| \text{ where } \begin{cases} p_{1} = \frac{2\sigma_{1}}{\cos\theta} \\ p_{2} = \frac{\sigma_{1} + \sigma_{2}}{\cos\theta} \\ p_{3} = \frac{2\sigma_{2}}{\cos\theta} \end{cases}$$

$$(8.44)$$

This has the advantage of having only four unknowns and six observables, as it assumes a relationship between the co- and cross-eigenpolarisation channels. This does, however, lead to an assumed rank ordering of the three polarisations that the Legendre approach of equation (8.43) does not require. Again, h_v and ϕ are common to all three channels, and the differences between polarisations are modelled by variation of extinction. There are no straight lines embedded in this equation, and solution is best tackled by a brute-force iterative search technique for the four-dimensional minimization.

8.2.6 OVOG model height estimation

In the second case to be considered we redirect attention to oriented volume problems where the influence of the underlying surface cannot be ignored, as



Fig. 8.21 Geometry of the oriented-volume-over-ground (OVOG) model

shown schematically in Figure 8.21. Here the underlying surface contributes a μ value in one or more polarisation channels, and so we cannot use the simple volume decorrelation models of the previous section. Again, by restricting attention to the co- and crosspolarised combinations of eigenpolarisations for layer 1, we can formally set up the following OVOG cost function:

$$G_{ovog} = \left\| \tilde{\gamma}_{xx} - e^{i\phi} (\tilde{\gamma}_{vo}^{xx} + F_{xx}(1 - \tilde{\gamma}_{vo}^{xx})) \right\| \\ + \left\| \tilde{\gamma}_{xy} - e^{i\phi} (\tilde{\gamma}_{vo}^{xy} + F_{xy}(1 - \tilde{\gamma}_{vo}^{xy})) \right\| \\ + \left\| \tilde{\gamma}_{yy} - e^{i\phi} (\tilde{\gamma}_{vo}^{yy} + F_{yy}(1 - \tilde{\gamma}_{vo}^{yy})) \right\|$$
(8.45)

This has six observables but seven unknowns, and hence is not well suited to solution as it stands. To be able to make inversion tractable we need to make some further assumptions about one or more of the parameters in the model. The simplest of these is the high μ dispersion assumption—also used in Section 8.1.4 for topography estimation. In this case we assume that the extinction in the *x* polarisation is so high as to reduce μ in this channel to zero, while the low extinction channel *y* maintains a high surface-to-volume scattering ratio. In this case we can simplify the cost function as shown in equation (8.46):

$$G_{ovog} = \|\tilde{\gamma}_{xx} - e^{i\phi}\tilde{\gamma}_{vo}^{xx}\| + \|\tilde{\gamma}_{xy} - e^{i\phi}(\tilde{\gamma}_{vo}^{xy} + F_{xy}(1 - \tilde{\gamma}_{vo}^{xy}))\| + \|\tilde{\gamma}_{yy} - e^{i\phi}(\tilde{\gamma}_{vo}^{yy} + F_{yy}(1 - \tilde{\gamma}_{vo}^{yy}))\|$$
(8.46)

This now has six unknowns and six observables. This can be further simplified if we use the xx and yy channels to estimate surface topography using the line fit technique, as described in Section 8.1.2. In this way we can reduce the balance of equation (8.46) to five unknowns and six observables. Problems arise with this approach if μ remains high in all channels. This can occur, for example, in applications involving thin layers with low to moderate extinctions at high angles of incidence. This causes several problems, as the coherences then all migrate down their lines to the topography point or find the true volume scattering channel for depth estimation. The only approach then is to constrain the range of extinctions expected in the problem (by physical modelling or external measurements), and then use these as known parameters in equation (8.46) to leave five unknowns for the six observables. Of course, if additional further information is available (for example, the phase ϕ of the surface from external measurements) then this can be added to further reduce the parameter imbalance. Whichever course is taken, the result is a set of inversions across the parameter range of extinctions. These provide us with a mean solution and error bars associated with the spread of solutions.

8.3 Hidden surface/target imaging

In this section we consider methods for using polarimetric interferometry to separate volume and surface contributions to the total backscattering crosssection of two-layer problems of the form shown in Figure 8.22. Here we show, on the left, the combined surface and volume scattering geometry, and on the right the main objective of isolating the (effective) surface components. Here there are two primary motives. The first is to be able to study the surface properties such as surface roughness and moisture, even in the presence of a vegetation or snow layer, for example (Cloude, 2005a). The second objective is to be able to image the surface beneath the volume (using synthetic aperture radar, for example) in order to detect objects located on the surface and obscured by the top layer (Sagues, 2001). This application includes foliage penetration, or FOPEN, in military detection as well as in search-and-rescue in forested or avalanche conditions (Cloude, 2004).

8.3.1 **RVOG** estimation of μ

The first step in this process is to identify a polarisation channel, which if possible contains only volume scattering and no surface component at all. We can again find the best approximation to such a channel by using physical arguments or the coherence optimization techniques of Section 6.2, where we identify the polarisation channel with smallest μ as the one with the largest phase bias. In either case this allows us to find a reference complex coherence for the volume scattering component. If we now assume that layer 1 is a random volume, then according to the SVOG model this point lies on a line joining the volume coherence to the surface phase point on the unit circle. If we now calculate the complex coherence in any other polarisation channel \underline{w} , then we find $F(\underline{w})$ —the fraction of surface scattering in this channel by a line fit between two complex values. This parameter can then be directly estimated from the



Fig. 8.22 Schematic representation of the hidden surface imaging problem or 2-to-1 layer conversion

two complex coherence values, as shown in equation (8.47):

$$F(\underline{w}) = \frac{-B - \sqrt{B^2 - 4AC}}{2A} \quad \mathbf{0} \le F(\underline{w}) \le 1$$

$$A = \left|\tilde{\gamma}_{\underline{w}_{\nu}}\right|^2 - 1 \quad B = 2\operatorname{Re}((\tilde{\gamma}(\underline{w}) - \tilde{\gamma}_{\underline{w}_{\nu}}).\tilde{\gamma}_{\underline{w}_{\nu}}^*) \quad C = \left|\tilde{\gamma}(\underline{w}) - \tilde{\gamma}_{\underline{w}_{\nu}}\right|^2 \right\} \Rightarrow \mu(\underline{w}) = \frac{F(\underline{w})}{1 - F(\underline{w})}$$

$$(8.47)$$

Note that this a just a special case of the general phase bias removal algorithm of equation (8.5), the main difference being that now the reference point is assumed to have $\mu = 0$ and F is the desired parameter, rather than just being an intermediate step towards phase estimation.

This approach works acceptably well for large μ , but for small μ the two coherences are close in the complex plane, and so any small errors in coherence estimation can lead to large errors in line fit. A more robust approach is therefore to find the estimated surface phase ϕ using large μ separations or a least squares line fit (as in equation (8.16)), and to then project all coherences onto the best-fit line joining the volume coherence to the estimated unit circle point *before* estimating *F* for a given polarisation \underline{w} , as shown in equation (8.48):

$$F(\underline{w}) = \frac{\tilde{\gamma}_p(\underline{w}) - \tilde{\gamma}_v}{e^{i\hat{\phi}} - \tilde{\gamma}_v}$$
(8.48)

Here the projected value of the coherence is given in terms of the line fit parameters m and c, as shown in equation (8.49):

$$x_{i} = \frac{\operatorname{Re}\left(\tilde{\gamma}(\underline{w})\right) + \hat{m}.\operatorname{Im}\left(\tilde{\gamma}(\underline{w})\right) - \hat{m}.\hat{c}}{1 + \hat{m}^{2}}$$
$$y_{i} = \hat{m}.x_{i} + \hat{c}$$
$$\left\{\tilde{\gamma}_{p}\left(\underline{w}\right) = x_{i} + iy_{i} \qquad (8.49)$$

This relation is derived by minimising the distance between the coherence $\tilde{\gamma}(\underline{w})$ and the line with known slope and estimating $y = \hat{m}x + \hat{c}$, posed as shown in equation (8.50):

$$\min_{x_i} \left\| \left(\operatorname{Re}(\tilde{\gamma}) - x_i)^2 + \left(\hat{m}x_i + c - \operatorname{Im}(\tilde{\gamma}) \right)^2 \right\|$$
(8.50)

This algorithm is summarized schematically in Figure 8.23, where we show the end points of the linear coherence region for the SVOG model assumption. The parameter *F* is then just the fractional distance along the line from the volume-only coherence point passing through the projected coherence towards the unit circle. Clearly, if the coherence approaches the unit circle F = 1 we have 100% surface scattering. This fractional parameter can then be directly used with an estimate of the total scattering cross-section σ to isolate the effective surface component, as shown in equation (8.51):

$$\sigma (\underline{w}) = \sigma_v (\underline{w}) + \sigma_{es} (\underline{w}) = (1 - F)\sigma (\underline{w}) + F\sigma (\underline{w})$$
$$\Rightarrow \sigma_{es} (\underline{w}) = G_s(\underline{w}) = F\sigma (\underline{w})$$
(8.51)

Note the following important points about this decomposition:

1) The effective surface component is attenuated by extinction through the volume, and hence is not the same as the bare surface return. However,



Fig. 8.23 Projection of general coherence onto the line model

under the SVOG assumption this attenuation is equal in all polarisation channels, and so polarisation ratios will be preserved. This is important, as several surface parameter retrieval algorithms employ ratios rather than absolute values. For example, surface moisture and roughness under the X-Bragg model (see Section 3.2.1) can be found from functions R and M as ratios of the Pauli scattering components. In the two-layer context we can now replace these formulae for bare surfaces with the following ratios, to be used to estimate parameters for a surface hidden beneath a random scattering layer:

$$R = \frac{G_s\left(\underline{w}_{HH-VV}\right) - G_s\left(\underline{w}_{HV}\right)}{G_s\left(\underline{w}_{HH-VV}\right) + G_s\left(\underline{w}_{HV}\right)} \quad M = \frac{G_s\left(\underline{w}_{HH-VV}\right) + G_s\left(\underline{w}_{HV}\right)}{G_s\left(\underline{w}_{HH+VV}\right)}$$
(8.52)

However, this result masks a problem with this approach: the confusion of direct and specular surface scattering in the polarimetric response, as we now consider.

2) The surface component in polarimetric interferometry is 'effective' in that it includes everything with a phase centre located on the surface. As we have seen in Section 7.3, this includes not just the direct surface return but also the specular second order scattering. Note that this acts to enhance $F(\underline{w}_{HH-VV})$ at the expense of $F(\underline{w}_{HH+VV})$ because of the π polarimetric phase change on specular reflection, and so will distort the moisture and roughness estimates of equation (8.52). A method of correcting for this is to estimate the full polarimetric coherency matrix for the surface components and use an incoherent decomposition to separate the direct and dihedral components *before* applying appropriate surface parameter techniques. One way to do this is to model the surface components as a rank-3 reflection symmetric coherency matrix, which can then be reconstructed from seven separate G_s estimates, as shown in equation (8.53):

$$[T_{s}] = m \begin{bmatrix} G_{s}(\underline{w}_{1}) & \frac{1}{2}(G_{s}(\underline{w}_{4}) - G_{s}(\underline{w}_{5}) - i(G_{s}(\underline{w}_{6}) - G_{s}(\underline{w}_{7}))) & 0\\ \frac{1}{2}(G_{s}(\underline{w}_{4}) - G_{s}(\underline{w}_{5}) + i(G_{s}(\underline{w}_{6}) - G_{s}(\underline{w}_{7}))) & G_{s}(\underline{w}_{2}) & 0\\ 0 & 0 & G_{s}(\underline{w}_{3}) \end{bmatrix}$$

where

$$\underline{w}_{1} = \begin{bmatrix} 1\\0\\0 \end{bmatrix} \quad \underline{w}_{2} = \begin{bmatrix} 0\\1\\0 \end{bmatrix} \quad \underline{w}_{2} = \begin{bmatrix} 0\\0\\1 \end{bmatrix} \quad \underline{w}_{4} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1\\1\\0 \end{bmatrix} \quad \underline{w}_{5} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1\\-1\\0 \end{bmatrix}$$

$$\underline{w}_{6} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1\\i\\0 \end{bmatrix} \quad \underline{w}_{7} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1\\-i\\0 \end{bmatrix}$$
(8.53)

We can then apply any of the incoherent decomposition theorems of Section 4.2 to separate the dihedral and direct components of this matrix.

We now turn to consider a sensitivity analysis of the various surface/volume separation algorithms. We have already seen that one way of estimating μ from

polarimetry alone (without the need for interferometry) is given by incoherent decomposition (see Section 4.2.3), the model-based form of which is summarized again in equation (8.54), where t_{ii} are the elements of the combined surface and volume coherency matrix and we set $F_p = 2$ (see equation (4.62)).

$$m_{v} = t_{33} \quad m_{d,s} = \frac{(t_{11} + t_{22} - 3t_{33}) \pm \sqrt{(t_{11} - t_{22} - t_{33})^{2} + 4|t_{12}|^{2}}}{2}$$
$$\alpha_{d,s} = \cos^{-1} \left[\left(1 + \left| \frac{t_{12}}{t_{22} - t_{33} - m_{d,s}} \right|^{2} \right)^{-\frac{1}{2}} \right]$$

 $m_{\max} = \max(m_d, m_s) \Rightarrow \alpha_{\max}$

$$\Rightarrow \mu_{\max} = \frac{m_{\max}}{m_v} (\sin^2 \alpha_{\max} + \frac{1}{2} \cos^2 \alpha_{\max})$$
(8.54)

Here we show how to estimate the maximum μ ratio from coherency matrix data. Key to this is an assumption about the depolarisation caused by the volume component, in that it has the characteristic diagonal 2:1:1 structure of dipole scattering. The problem with this approach is that the volume term has a high scattering entropy and hence requires a large number of data samples to reduce the variance of the estimated coherency matrix to sufficiently low levels to be able to isolate any small surface contributions (Lopez-Martinez, 2005). Furthermore, this speckle fluctuation can also lead to negative estimates for scattered powers m_d and m_s , unless explicit attempts are made to enforce the positive semi-definite nature of [T].

As a measure of this lack of sensitivity we consider, as an example, a simplified problem with a volume-only contribution— $\mu_{max} = 0$ and T = diag(2,1,1)—and then use the Monte Carlo technique of Appendix 3 to generate random samples taken from a normal distribution with the same underlying coherency matrix. We then use these samples to estimate the mean μ_{max} as a function of increasing number of samples. In the limit of an infinite number of samples we will of course obtain $\mu_{max} = 0$, but we see in Figure 8.24 that the convergence is rather slow, with more than 100 samples required to be able to identify -10 dB of surface contribution in a volume scattering background. Again, this can be traced to the high scattering entropy of the volume. On the other hand, coherent methods based on polarimetric interferometry are potentially more sensitive, as the volume decorrelation is now a function of the baseline/height product, which can be designed to optimize performance. In addition, they involve relaxed assumptions about the nature of the volume scattering (as long as it maintains azimuthal symmetry).

8.3.2 Optimum baseline for hidden surface detection

In contrast, we can express the sensitivity of interferometric coherence to surface effects caused by μ by calculating the fractional change in the length of the coherence line due to the presence of a surface component, as shown in Figure 8.25. Here we can see that even with $\mu = -10$ dB, the shift in coherence is around 10% of the total line length. Therefore, in order to be able to detect a small change in μ we need to choose the baseline so that the corresponding



Fig. 8.24 Estimation of apparent surface scattering contribution in volume-only scattering for the Freeman-eigenvalue model versus number of looks

Fig. 8.25 Fractional line length in SVOG model (*F*) versus relative level of surface scattering (μ)

change is detectable. To do this we therefore design the interferometer to have a long line length in the complex plane, so we can be sensitive to the presence of small surface components. The total line length itself is just given by $|1 - \tilde{\gamma}_{vo}|$, and hence determination of line length involves assumptions about the volume-only coherence $\tilde{\gamma}_{VO}$. On the other hand, we also wish to minimize the number of data samples (*L*) required to ensure accurate estimates. These two requirements are in conflict, and require some compromize through correct baseline selection, as follows. We start by calculating the derivative of coherence with

respect to the surface-to-volume ratio μ , as shown in equation (8.55):

$$\frac{\partial \hat{\gamma}}{\partial \mu} = \frac{(1 - \tilde{\gamma}_{vo})}{(1 + \mu)^2} = f(\mu)g\left(\tilde{\gamma}_{vo}\right)$$
(8.55)

Maximising sensitivity would then seem to require that $\gamma_{vo} = -1$; that is, the baseline is chosen so that the phase centre for volume scattering lies at the π height of the interferometer. To be realized, however, this phase must also occur with a coherence magnitude of unity. This is not a realistic scenario, requiring as it does infinite extinction in RVOG or a very localized structure function in the Legendre approximation. Adopting the latter, we can express the line length more realistically as a function of three parameters, $k_v a_{20}$ and a_{10} , as shown in equation (8.56):

$$|1 - \tilde{\gamma}_{\nu o}| \approx \left| 1 - e^{ik_{\nu}} (f_0 + a_{10}f_1 + a_{20}f_2) \right|$$
(8.56)

Before considering this in more detail, we first include the change of *minimum* coherence along the line. This is important, as it impacts on the number of samples required to estimate coherence and hence on the accuracy and resolution of any estimation. The μ for minimum coherence was found in equation (7.49). Inserting this into the line coherence model we obtain the following expression for the minimum coherence:

$$\begin{split} \tilde{\gamma}(\underline{w}) &= e^{i\phi(z_o)}(\tilde{\gamma}_{vo} + F(\underline{w}) (1 - \tilde{\gamma}_{vo})) \\ \Rightarrow L &= \left| \tilde{\gamma}(\underline{w}) \right|^2 = a + bF + cF^2 \Rightarrow \frac{dL}{dF} = b + 2cF \\ \Rightarrow F_{\min} &= -\frac{b}{2c} = \frac{|\tilde{\gamma}_{vo}|^2 - \operatorname{Re}(\tilde{\gamma}_{vo})}{(1 - \tilde{\gamma}_{vo}) (1 - \tilde{\gamma}_{vo}^*)} \\ \Rightarrow |\tilde{\gamma}_{\min}| &= \left| \tilde{\gamma}_{vo} + \frac{|\tilde{\gamma}_{vo}|^2 - \operatorname{Re}(\tilde{\gamma}_{vo})}{(1 - \tilde{\gamma}_{vo}) (1 - \tilde{\gamma}_{vo}^*)} (1 - \tilde{\gamma}_{vo}) \right| \\ &= \left| \frac{\operatorname{Im}(\tilde{\gamma}_{vo})}{(1 - \tilde{\gamma}_{vo}^*)} \right| \end{split}$$
(8.57)

The requirement of keeping this minimum as high as possible is in conflict with the simultaneous desire to maximize the line length of equation (8.56). As a compromise we choose to select a k_v value that maximizes the product of equations (8.56) and (8.57); that is, that maximizes the expression in equation (8.58):

$$|1 - \tilde{\gamma}_{vo}| \cdot |\tilde{\gamma}_{\min}| = \left| \frac{\mathrm{Im}(e^{ik_v}(f_0 + a_{10}f_1 + a_{20}f_2))}{1 - e^{-ik_v}(f_0 - a_{10}f_1 + a_{20}f_2)} \right| \cdot \left| 1 - e^{ik_v}(f_0 + a_{10}f_1 + a_{20}f_2) \right|$$
(8.58)

We can now investigate the upper and lower bounds of this function as we change structure for a given baseline/height product. Note that a_{10} and a_{20} have a limited range, as we are considering the volume-only component of the structure function (and so only structure functions that increase with height).



Fig. 8.26 Bounds on line length versus k_v for all structure parameters in the range $a_{10} \ge 0, a_{20} \le 0$

Fig. 8.27 Bounds on minimum coherence as a function of k_v for all structure parameters in the range $a_{10} \ge 0, a_{20} \le 0$

This limits the Legendre spectrum, so that a_{10} is non-negative and a_{20} is nonpositive. In Figures 8.26 and 8.27 we show how the two components of this function vary with k_v . The line length in equation (8.26) starts at zero, and then increases to a maximum of 2 before falling again for high k_v . However, at the same time we see the minimum coherence value start at 1 for low k_v , and decrease to zero when the line goes through the origin. Indeed, we see that the minimum can be zero for all structure configurations beyond $k_v = 2$.

Finally, in Figure 8.28 we show the corresponding variation of the product of these two components, and see a clear optimum range $1 \le k_v \le 1.5$. This range then represents the best compromise between line-length (sensitivity) and



Fig. 8.28 Bounds on product of line length/ minimum coherence versus k_v for all structure parameters in the range $a_{10} \ge 0, a_{20} \le 0$

number of samples required for estimation (resolution). We see that a design value around $k_v = 1.25$ (centre of the range) represents a good choice. We can then calculate the optimum baseline to be used by first specifying a target layer height h_{design} . This can then be used with the baseline geometry (see Section 5.1) to calculate the spatial baseline required, *B*, for wavelength λ , as shown in equation (8.59):

$$k_{\nu} = \frac{\beta_z h_{design}}{2} = \frac{2\pi B_n h_{design}}{\lambda R \sin \theta} = 1.25 \Rightarrow B = \frac{1.25\lambda R \sin \theta}{2\pi h_{design} \cos \left(\theta - \delta_d\right)}$$
(8.59)

Note that the coherence so obtained can be much higher than the polarimetric only coherence, and so by adopting interferometry we obtain a better situation (effectively a lower entropy) than the polarimetric approach based on equation (8.54).

8.4 Structure estimation: extinction and Legendre parameters

In this section we consder methods for estimating the vertical structure function f(z) itself; that is, of estimating the vertical variation of scattering through layer 1 using polarimetric interferometry. This information is useful for classification of different layer types (in forestry and vegetation, for example, where canopy depth can be an indicator of species or plant stress), and for the estimation of propagation parameters such as the mean total and differential wave extinction, from which we can then indirectly obtain information about water content and density (Ballester-Berman, 2005; Lopez-Sanchez, 2006, 2007; Cloude, 2006b, 2007a).

In the RVOG and OVOG models, polarisations other than volume-only are related by a scale factor: the surface-to-volume scattering ratio, μ . This parameter can be estimated for an arbitrary polarisation using the techniques described in equation (8.48). The corresponding vertical profile is then obtained as a weighted sum of an exponential and delta function, as shown in equation (8.60):

$$\hat{f}_{rvog}(\underline{w}, z) = \frac{\cos\theta e^{-\frac{2\theta_c}{\cos\theta}z}}{2\hat{\sigma}_e(e^{\frac{2\theta_c}{\cos\theta}\hat{h}_v} - 1)} + \hat{\mu}(\underline{w})\delta(z) \quad 0 \le z \le \hat{h}_v$$
(8.60)

In the Legendre approach, more interesting possibilities arise and lead to a generalization of the structure estimation problem, termed coherence tomography (Cloude, 2006b), as we now consider.

8.4.1 Coherence tomography (CT)

Our starting point is to adopt the second-order Legendre expansion of coherence as shown in equation (8.30). This allows us to include quadratic as well as linear and constant scattering profiles. Here we have four unknowns on the right: k_v , ϕ_0 , and the two normalized Legendre coefficients a_{10} and a_{20} . On the left we have only two observables (the complex coherence). We note that only two of the parameters depend on polarisation. The next stage is therefore to isolate the polarisation-dependent terms, as shown in equation (8.61), where a_{00} has unit value, and where the real functions f_0 , f_2 and imaginary function f_1 as defined in equation (5.55), are independent of polarisation and depend only on k_v .

$$\tilde{\gamma}(\underline{w})e^{-i(k_v+\phi_o)} = \tilde{\gamma}_k \approx a_{00}f_0 + a_{10}\left(\underline{w}\right)f_1 + a_{20}\left(\underline{w}\right)f_2 \tag{8.61}$$

This can then be written in matrix form in terms of the real and imaginary parts of the phase normalized coherence, as shown in equation (8.62):

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & -if_1 & 0 \\ 0 & 0 & f_2 \end{bmatrix} \cdot \begin{bmatrix} a_{00} \\ a_{01} \\ a_{02} \end{bmatrix} = \begin{bmatrix} 1 \\ \operatorname{Im}(\tilde{\gamma}_k) \\ \operatorname{Re}(\tilde{\gamma}_k) - f_0 \end{bmatrix} \Rightarrow [L]\underline{a} = \underline{b} \quad (8.62)$$

The next important idea is that we can now invert this relationship to obtain estimates of the polarisation-dependent Legendre parameters from coherence and knowledge of the matrix [L], as shown in equation (8.63):

$$\underline{\hat{a}} = [L]^{-1} \underline{\hat{b}} \tag{8.63}$$

From the vector \hat{a} we can then estimate the normalized vertical structure function for a known layer depth h_v , as shown in equation (8.64):

$$0 \le z \le h_{\nu} \Rightarrow \hat{f}_{L2}(z) = \frac{1}{h_{\nu}} \left\{ 1 - \hat{a}_{10} + \hat{a}_{20} + \frac{2z}{h_{\nu}} (\hat{a}_{10} - 3\hat{a}_{20}) + \hat{a}_{20} \frac{6z^2}{h_{\nu}^2} \right\}$$
(8.64)

Note that when the quadratic term (a_{20}) is zero this reverts to the linear approximation $\hat{f}_{L1}(z)$, as developed in equation (8.27). Equations (8.62) and (8.63) constitute a method for reconstructing the function f(z) from coherence, and

are therefore termed coherence tomography, or CT. We shall see that the matrix formulation can be extended to arbitrary order of Legendre polynomial, and hence to higher and higher resolution reconstructions by adding multiple baselines to the interferometer. However, there remains the important issue of how to obtain estimates of the polarisation-independent terms k_v and ϕ_0 . We now turn to consider this in more detail.

In equation (8.61) we separated coherence into polarisation-dependent and independent components. The latter set comprises three parameters of interest: the layer depth h_v and interferometric wavenumber β_z (which are then used to calculate k_v), and the phase of the bottom of the layer, ϕ_0 . The wavenumber can be estimated from knowledge of the baseline geometry and operating wavelength of the interferometer (see Section 5.1), but the other two parameters require special attention. There are two principle ways of obtaining these parameters.

8.4.1.1 CT using external data

In the first approach we can use separate external measurement of the layer depth h_v and surface phase (the latter by measuring the *z* coordinate z_0 of the bottom of the layer above the zero datum of the interferometer, and using β_z to obtain $\phi_0 = \beta_z z_0$). These can be obtained, for example, for laboratory-based experiments (Cloude, 2007a) by direct measurement, and then directly used in equation (8.60) to investigate the variation of structure function with polarisation. In field experiments such estimation can be more difficult, but can still be accomplished with the aid of global positioning technology such as GPS, or depth profiling technologies such as laser sounding using LIDAR or high-resolution microwave altimeters or scatterometers.

In this case we can estimate the structure function for arbitrary polarisations \underline{w} by first forming the interferogram, estimating complex coherence, phase shifting the coherence using β_z and ϕ_0 , and then calculating the profile estimate as summarized in equation (8.65):

$$\begin{cases} \beta_{z} \\ h_{v} \\ \phi_{0} \end{cases} \rightarrow k_{v} = \frac{\beta_{z}h_{v}}{2}, \quad \tilde{\gamma}_{k}(\underline{w}) = \tilde{\gamma}(\underline{w})e^{-i(k_{v}+\phi_{0})} \\ f_{o} = \frac{\sin k_{v}}{k_{v}} \quad f_{1} = i\left(\frac{\sin k_{v}}{k_{v}^{2}} - \frac{\cos k_{v}}{k_{v}}\right) \quad f_{2} = \frac{3\cos k_{v}}{k_{v}^{2}} - \left(\frac{6 - 3k_{v}^{2}}{2k_{v}^{3}} + \frac{1}{2k_{v}}\right)\sin k_{v} \\ \begin{bmatrix} 1 & 0 & 0 \\ 0 & -if_{1} & 0 \\ 0 & 0 & f_{2} \end{bmatrix} \cdot \begin{bmatrix} a_{00} \\ a_{01}(\underline{w}) \\ a_{02}(\underline{w}) \end{bmatrix} = \begin{bmatrix} 1 \\ \operatorname{Im}(\tilde{\gamma}_{k}(\underline{w})) \\ \operatorname{Re}(\tilde{\gamma}_{k}(\underline{w})) - f_{0} \end{bmatrix} \\ \Rightarrow \hat{g}_{L2}(\underline{w}, z) = \frac{1}{h_{v}} \left\{ 1 - \hat{a}_{10}(+\hat{a}_{20}(\underline{w}) + \frac{2z}{h_{v}}(\hat{a}_{10}(\underline{w}) - 3\hat{a}_{20}(\underline{w})) + \hat{a}_{20}(\underline{w}) \frac{6z^{2}}{h_{v}^{2}} \right\}$$

$$(8.65)$$

This approach makes no assumptions about the shape of the coherence region, and so can be used to investigate the most general profiles. We will consider an example of such an approach based on laboratory anechoic chamber measurements of maize plants in Chapter 9. Often, however, especially in remote sensing applications, we have no access to the layer depth or supporting measurements, and must therefore develop alternative techniques for estimating the parameters directly from the data itself. There are two approaches to be considered: dual baseline inversion, when we add new baselines to increase the number of observables, and single baseline bootstrap techniques, where we use the height and surface phase estimators of Sections 8.1 and 8.2 to enable tomography with a single baseline. We now turn to consider such methods in more detail.

8.4.1.2 Dual baseline inversion

The challenge is now to develop parameter estimation algorithms for the surface phase ϕ_0 and layer depth h_v that involve minimal assumptions about the shape of the structure function f(z). In this way we can use these estimates in the second-order Legendre algorithm directly, and maintain its flexibility to deal with general scattering scenarios.

The algorithms presented in Sections 8.1 and 8.2 for ϕ_0 and h_v estimation can be proposed, but all of them made some further restrictive assumptions about f(z) in the volume-only channel—that it is exponential for the RVOG and OVOG models, or that the volume-only scattering coefficients satisfy $a_{20} \leq 0$ and $a_{20} \geq 0$ for the Legendre approach. These assumptions strictly only have to be valid for the volume-only polarisation channel, but because of the SVOG assumption used to estimate topography they impact on the assumed volume component of the response in arbitrary polarisation channels. Hence such assumptions force the reconstructions to conform to a subset of structure functions satisfying the requirements of the models. Only if these assumptions are a good match to the physical structure of the problem will they yield good results. It is therefore of interest to see if we can avoid such restrictive assumptions at all.

Here we consider one important way to achieve this, by using a dual-baseline interferometer, with a second baseline, different from the first, used to obtain four observables (the amplitude and phase of two coherences) with four model unknowns (h_v , ϕ_0 , a_{10} , and a_{20}). We start with the simpler case when ϕ_0 is known for both baselines, and so we can set $\phi_0 = 0$ without loss of generality. We shall consider the full four-dimensional case later. This then reduces the problem to three unknowns (h_v , a_{10} , and a_{20}). Even so, as we have seen, problems arise with CT if we do not know the k_v value in advance. In this case we obtain multiple solutions for a whole range of $k_v a_{10}$, a_{20} coordinates. As shown in Figure 8.29, a single coherence point (in grey) can fit the model over



Fig. 8.29 Superimposed second-order Legendre approximation for two different k_v values, showing structural ambiguity for a single-baseline coherence (in grey)



Fig. 8.30 Schematic representation of wellconditioned (left) and ill-conditioned (right) solutions for dual-baseline inversion

a wide range of k_v values. Here we show the set of coordinate lines for two k_v values, with the two origins ($a_{10} = a_{20} = 0$) shown as black points, and see that the sample coherence point, although it has different $a_{10}a_{20}$ values, can be made to fit either. Hence a single baseline cannot be used to estimate the two Legendre structure parameters uniquely. We can, however, estimate a family of solutions using single baseline data. As we move around the SINC spiral in Figure 8.29 we obtain a set of solution pairs a_{10}, a_{20} for the given coherence (in grey). We can represent this family geometrically as a set of solution points in the a_{10}, a_{20} plane, generated as k_v varies from 0 to π . These generate a curve in the plane, as shown schematically in Figure 8.30. We can then use this idea to propose a method for estimating the true k_v value by combining data from a second baseline.

If we now consider that coherence data is available for a second additional baseline, related to the first by a baseline ratio B_r , then we can generate a second set of a_{10} , a_{20} points by simultaneously solving the matrix equation for k_v for the first baseline and $B_r k_v$ for the second. The correct k_v value then occurs when these two curves intersect, as shown schematically in Figure 8.30. Here we show two possible scenarios. On the left, the solid curve is the curve of solutions obtained for the first baseline, and the dashed curve is the solutions for the second. This is a well-conditioned case, when the intersection point occurs for nearly orthogonal curves. Such a scenario will be robust to errors in the two coherence estimates. On the right of the figure is shown the opposite case of a poorly conditioned solution where the intersection point occurs for nearly parallel curves. For example, in the limiting case, if we take the two baselines equal $(B_r = 1)$ then the loci will exactly overlap and a solution is not possible. When nearly overlapping, any small perturbation of the curves will lead to a large change in the solution. This ill-conditioning can undermine the uniqueness of a solution using dual baselines by making the algorithm so sensitive to noise that it cannot be used in practical applications (Hopcraft, 1992). The level of such ill-conditioning will be a function of the baseline ratio B_r . We shall examine the conditioning of coherence tomography and its dependence on B_r in more detail in Section 8.4.3. First, however, we consider a formalization of this approach and how to generalize it for unknown surface topography ϕ_0 .

We can formally write the solution to the dual baseline k_v estimate as minimization of the coherence error as defined in equation (8.66), where subscripts

1 and 2 refer to the baselines used:

$$\begin{split} \tilde{\gamma}_{1} &= e^{i\phi_{1}}e^{ik_{v1}}\left(f_{0}(k_{v1}) + a_{10}f_{1}(k_{v1}) + a_{20}f_{2}(k_{v1})\right) \Rightarrow \tilde{\gamma}_{k1} = \tilde{\gamma}_{1}e^{-i\phi_{1}}e^{-ik_{v1}}\\ \begin{bmatrix} \hat{a}_{00}\\ \hat{a}_{01}\\ \hat{a}_{20} \end{bmatrix} &= \begin{bmatrix} 1 & 0 & 0\\ 0 & -if_{1}(k_{v1}) & 0\\ 0 & 0 & f_{2}(k_{v1}) \end{bmatrix}^{-1} \cdot \begin{bmatrix} 1\\ \mathrm{Im}(\tilde{\gamma}_{k1})\\ \mathrm{Re}(\tilde{\gamma}_{k1}) - f_{0}(k_{v1}) \end{bmatrix}\\ \Rightarrow \tilde{\gamma}_{2}^{est} &= e^{iB_{r}\phi_{1}}e^{ik_{v2}}\left(f_{0}(k_{v2}) + \hat{a}_{10}f_{1}(k_{v2}) + \hat{a}_{20}f_{2}(k_{v2})\right) \text{ where } k_{v2} = B_{r}k_{v1}\\ \Rightarrow \text{ Coherence error} &= \|\tilde{\gamma}_{2} - \tilde{\gamma}_{2}^{est}\| \end{split}$$
(8.66)

The procedure (for $\phi_1 = 0$) is then to vary k_{v1} from 0 to π , calculating for each value the Legendre spectrum a_{10}, a_{20} . We then use these values to *estimate* the second baseline coherence and select the triplet k_{v1}, a_{10}, a_{20} that minimizes the difference between this complex estimate and the true second baseline coherence $\tilde{\gamma}_2$, as shown in equation (8.66). When topographic phase is also unknown, the only difference we face is to search for the intersection of a_{10}, a_{20} loci in a two-dimensional space of ϕ_0 and k_{v1} rather than just k_{v1} . The estimate of the second baseline coherence is then phase shifted by a scaled topography, as shown in equation (8.66). (Note that we are also assuming that there are no residual phase errors between baselines, as can occur, for example, in repeat pass sensors. If not true, then we must add an extra unknown phase parameter to equation (8.66).)

As a typical example of dual baseline performance, consider the choice ($\phi_1 = \phi_2 = 0$) $B_r = 0.5$ and $k_v = \pi/2$, with a profile defined by $a_{10} = 0.5$ and $a_{20} = 1$. These lead to very distinct coherences of 0.87 for the smaller baseline and 0.54 for the larger. These two points can then be used to estimate a pair of solution curves for a_{10} , a_{20} , as shown in Figure 8.31. Here we see a scenario that seems poorly conditioned, despite the fact that the coherences from the two baselines are very different, with the intersection point of the two curves occurring for



Fig. 8.31 Example solution loci for dualbaseline inversion, showing ill-conditioned nature of solution



Fig. 8.32 Variation of coherence error for dual-baseline inversion example of Figure 8.31, showing a minimum at correct value $(\pi/2)$

Fig. 8.33 Coherence error for twodimensional search in baseline/height product k_{v1} and surface phase f_1 (for true values of $k_v = \pi/2$ and $f_1 = \pi$)

nearly parallel sections. In Figure 8.32 we show the corresponding coherence error of equation (8.66), which correctly shows a unique minimum for $k_v = \pi/2$. Turning now to the general case when topography is also unknown, we consider a situation where $\phi_1 = \pi, \phi_2 = \pi/2$. Figure 8.33 shows the two-dimensional variation of coherence error. Again, formally, the correct solution is located with a minimum at $k_v = \pi/2, \phi_1 = \pi$, but we see a long 'valley' of potential local minima stretching across the solution space.

The level of ill-conditioning is especially important when we consider that the longer baseline coherence is quite low (around 0.54), and so will require a large

number of samples to minimize residual estimation noise. Such noise could lead to large errors in the estimate of k_v , and we need to balance the amplification of error caused by the ill-conditioning against the number of samples required in order to correctly assess these errors. This example nicely illustrates how uniqueness is not the only criterion required for an assessment of algorithm performance, and that the level of numerical stability or ill-conditioning must also be quantified. We will consider such an analysis in Section 8.4.3.

Equation (8.66) represents one method of estimating k_v , but it requires data for two baselines. But this is often not available, and so it is of interest to consider alternative single baseline strategies for estimation of k_v that still enable use of the second-order Legendre approach to CT.

8.4.2 Bootstrap polarisation coherence tomography (PCT)

In this approach we try to use single baseline data itself to approximate the two parameters ϕ_0 and k_v . The easiest case to deal with is the estimation of topography ϕ_0 . Here, by assuming only validity of the SVOG model—that layer 1 shows azimuthal scattering symmetry and is random so that polarisation dependence of coherence comes only from the variations of surface-to-volume scattering ratio—we again obtain a linear coherence loci, and can therefore use any of the line fit techniques developed in Section 8.1 to estimate topographic phase. Note that this symmetry makes no assumptions about the shape of the volume-only structure function, and assumes only that it is invariant in shape (but not necessarily in amplitude) to polarisation.

For k_v estimation using a single baseline configuration we must employ an algorithm for layer depth estimation that is robust to changes in structure and hence robust to changes in the a_{10} , a_{20} coefficients. We have already seen an example of such an algorithm in equation (8.38), where we used separate phase and coherence estimates to balance the errors across a range of structure parameters.

With this established, we can provide a direct estimate of k_v by first identifying a volume-dominated polarisation channel and then calculating k_v directly, as shown in equation (8.67):

$$k_{\nu} = \frac{1}{2} \left\{ \arg(\tilde{\gamma}_{w_{\nu}} e^{-i\phi_0}) + 0.8(\pi - 2\sin^{-1}(\left|\tilde{\gamma}_{w_{\nu}}\right|^{0.8}) \right\}$$
(8.67)

The estimated ϕ_0 , k_v values then establish a unique coordinate system for the whole coherence diagram, so that the structure function for all other polarisations can be reconstructed up to second order using the matrix inversion of coherence tomography.

This approximation works with only a single baseline, but requires the combination of at least two interferograms formed for different polarisations: one volume and the other surface dominated. It therefore requires some polarisation diversity in measurements, and hence we term it polarisation coherence tomography, or PCT, to distinguish it from standard CT. The basic steps involved in PCT are summarized in Figure 8.34. Here we show how, starting from a pair of polarisations, w_s and w_v , we can use the SVOG assumption to fit a line

Stage 1: Height and phase estimation



Fig. 8.34 Summary of algorithm for singlebaseline polarization coherence tomography (PCT)

and find the topographic phase. (The two-state line fit is shown in Figure 8.34, although the total least squares or other approach could equally be applied if more polarisation channels were to be made available.) We must then select one of the polarisations as a volume-dominated channel to obtain an estimate of k_v . We then use the k_v and ϕ_0 estimates to scale the coherence for arbitrary polarisation \underline{w} to generate a second order estimate of the structure function, as shown. This can then be repeated for arbitrary polarisations as desired.

The procedure has a simple geometric interpretation as shown in Figure 8.35. The process begins by identifying the volume coherence (grey point) and correcting for surface topography (white point) using the SVOG linear coherence assumption (black line). From the grey volume coherence we then estimate k_v , which locates the origin for the a_{10} , a_{20} coordinate system (the black point in Figure 8.35). This origin then defines a grid of $a_{10}a_{20}$ values that can be used to estimate the structure function for arbitrary location in the complex unit circle. However, under the SVOG assumption, coherence varies with polarisation





along the black line and so we can access only a limited portion of the a_{10}/a_{20} space. Note, for example, that while a_{10} can be positive or negative, the line permits only positive a_{20} values in the reconstruction. This is the price to pay for using such a bootstrap approach.

We now turn to consider more quantitatively the ill-conditioning of the matrix inversion embedded in both CT and PCT algorithms.

8.4.3 Condition number and error analysis

We have seen that estimation of the structure function using CT and PCT involves the following key matrix inversion step:

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & -if_1 & 0 \\ 0 & 0 & f_2 \end{bmatrix} \cdot \begin{bmatrix} a_{00} \\ a_{01} \\ a_{02} \end{bmatrix} = \begin{bmatrix} 1 \\ \operatorname{Im}(\tilde{\gamma}_k) \\ \operatorname{Re}(\tilde{\gamma}_k) - f_0 \end{bmatrix} \Rightarrow [L]\underline{a} = \underline{b} \Rightarrow \underline{\hat{a}} = [L]^{-1} \underline{b}$$
(8.68)

where the functions f_0 , f_1 and f_2 are given in Figure 8.34. Some care must be taken with this inversion, as it can lead to an amplification of any errors in <u>b</u> so that the resulting <u>a</u> may represent a very poor estimate of structure. An alternative way of thinking about this is to estimate the level of noise we can tolerate in the coherence vector <u>b</u> so as to keep the fractional error in <u>a</u> below a prescribed value. This will then allow us to estimate the number of coherence samples required for good estimation. In this section we develop such an algorithm for analysing the stability of matrix inversions like equation (8.68).

Key to quantifying this amplification process is the condition number (CN) of the matrix [L] (see Appendix 1). The larger CN, the larger any amplification of errors in \underline{b} . As [L] is diagonal in equation (8.68), we can obtain an explicit expression for the condition number of the matrix [F] as a ratio of the functions



Fig. 8.36 Variation of matrix condition number versus kv for single-baseline coherence tomography

 f_i , as shown in equation (8.69):

$$CN = -\frac{1}{f_2} = -\frac{k_\nu^2}{3\cos k_\nu - (3 - k_\nu^2)\frac{\sin k_\nu}{k}}$$
(8.69)

Figure 8.36 plots this function versus normalized wavenumber k_v . Note that for small baseline/height products the inversion is very poorly conditioned (a large CN). For baseline/height products around unity, the condition number is around 10–20. Since [*F*] is diagonal we can also identify the worst-case scenario, when the system becomes most sensitive to errors in <u>b</u>. From equation (8.68) this arises for perturbations of a true solution of the form

$$\underline{b} = \begin{bmatrix} 1\\0\\0 \end{bmatrix} \Rightarrow \underline{b} + \delta \underline{b} = \begin{bmatrix} 1\\0\\\delta \end{bmatrix}$$
(8.70)

This physically corresponds to small radial coherence amplitude perturbations about uniform zero-extinction volume scattering. In this worst case, the error in the Legendre coefficient vector is amplified by the matrix inversion to the order of CN. β . The coefficient β can now be related to the coherence and effective number of looks *L* by using the Cramer–Rao bound (see Appendix 3) and considering the limiting case of zero-extinction volume scattering. Considering the worst case from equation (8.70), it follows that the largest error contribution is from the real part of the phase corrected coherence $\tilde{\gamma}_k$. For a uniform zero-extinction volume the real part error is then dominated by the Cramer– Rao variance on coherence rather than phase estimation. Taking the standard deviation as a measure of the coherence error we can then write:

$$\beta \approx \frac{(1-\gamma_{\nu}^2)}{\sqrt{2L}} \approx \frac{1}{\sqrt{2L}} \left(1 - \left(\frac{\sin k_{\nu}}{k_{\nu}}\right)^2 \right)$$
(8.71)



Fig. 8.37 Maximum bound on fractional error in Legendre estimates versus *kv* for different number of looks

where *L* is the number of looks. In this way we can estimate an upper bound on the fractional error in the estimate of the Legendre coefficients as a function of just two parameters, k_v and the number of looks *L*, as shown in equation (8.72):

$$\max\left(\frac{\|\underline{\delta}\underline{a}\|}{\|\underline{a}\|}\right) = CN.\beta = \frac{\sin^2 k_v - k_v^2}{\sqrt{2L}\left(3\cos k_v - (3 - k_v^2)\frac{\sin k_v}{k_v}\right)}$$
(8.72)

Figure 8.37 shows how this bound varies as a function of k_v for various number of looks *L*. We should note that this represents a worst-case scenario, and generally the errors will be better than this. This approach assumes that the layer is almost a uniform volume scatterer (and so the volume coherence lies along the bounding SINC curve in Figure 8.35). If this is not true, and the volume channel has some other structure, then the errors will be less than this bound.

This conditioning error is due to amplified noise in coherence estimation. However, we have two other main sources of system noise to consider: the effects of SNR, and temporal decorrelation in the interferometer. These can now be incorporated into the CT and PCT formulations as follows.

8.4.4 SNR and temporal decorrelation in CT

Signal-to-noise ratio and temporal decorrelation effects can be included in the CT formalism by noting that they act as scalar multiplying factors of the observed coherence (see Section 5.2.5). Hence they do not distort the mean phase of the complex coherence but reduce the coherence amplitude (increase phase variance). This then scales the real and imaginary parts of the <u>b</u> vector as shown in equation (8.73). Note that they do not influence f_0 , which has now



Fig. 8.38 Effect of temporal decorrelation on the RVOG/second-order Legendre combination

been separated in the formulation.

$$\begin{bmatrix} a_{00} \\ a_{10} \\ a_{20} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -if_1 & 0 \\ 0 & 0 & f_2 \end{bmatrix}^{-1} \begin{bmatrix} 1 & 0 & 0 \\ 0 & \gamma_{snr}\gamma_t & 0 \\ 0 & 0 & \gamma_{snr}\gamma_t \end{bmatrix}^{-1} \begin{bmatrix} 1 \\ \operatorname{Im}(\tilde{\gamma}_k) \\ \operatorname{Re}(\tilde{\gamma}_k) \end{bmatrix} - \begin{bmatrix} 1 & 0 & 0 \\ 0 & -if_1 & 0 \\ 0 & 0 & f_2 \end{bmatrix}^{-1} \begin{bmatrix} 0 \\ 0 \\ f_0 \\ g. \\ (8.73) \end{bmatrix}$$

These have the geometrical effect of shifting the coherence point along a radial line towards the origin, as shown schematically for the volume coherence (grey point) in Figure 8.38. In CT, the effect will be to amplify the quadratic component of the structure function (with an increased positive a_{20} coordinate) without influencing the [L] matrix elements. For bootstrap PCT, however, we use the volume coherence itself to estimate k_v , and so temporal/SNR decorrelation will impact on the [L] matrix as well as the <u>b</u> vector. We see from Figure 8.38 that a radial shift will initially cause an overestimation of k_v . This will continue until the radial line intersects the SINC locus (shown as the light grey point in Figure 8.38).

For larger SNR/temporal decorrelation the volume coherence moves below the SINC boundary, and the Legendre approximation no longer provides a solution for k_v . We see from the curvature of the SINC locus that such SNR and temporal effects (which are independent of baseline) will be more serious for small k_v , where the locus approaches the unit circle. Therefore, one way to provide increased robustness to such effects is to work at larger spatial baselines. The best way to minimize such effects, however, is to avoid temporal effects by employing a single-pass interferometer and to ensure high SNR in the selected polarisation channels.

8.4.5 Multiple baseline CT

In the previous sections we saw how we can employ first- and second-order Legendre approximations of the structure function to model complex coherence. These relationships, under certain assumptions, can be inverted to provide estimates of the structure function from coherence measurements in a technique called coherence tomography (CT). However, such reconstructions are limited, as we have seen, to second-order polynomial variation. It is natural to ask if we can further improve the resolution of the reconstruction by estimating higher-order terms of the Fourier–Legendre expansion. In this section we consider such an extension, and show how, with knowledge of layer depth and surface position, we can employ multiple baseline interferometry to reconstruct the structure function to higher and higher resolutions—albeit at the price of increasing condition number with increasing resolution (Cloude, 2007a).

In single-baseline CT we have two observables (one complex coherence) and two unknowns (a_{10} and a_{20}), assuming we have knowledge of layer depth and surface position. Hence the addition of a second baseline adds two new observables and allows us to further extend the Legendre series by a further two orders to fourth order, as shown in equation (8.74). The new functions f_3 and f_4 are given in equation (8.75), where we note that f_3 is pure imaginary and f_4 is real:

$$\tilde{\gamma} e^{-ik_v} e^{-i\phi_o} = \tilde{\gamma}_k = f_0 + a_{10}f_1 + a_{20}f_2 + a_{30}f_3 + a_{40}f_4$$
(8.74)

$$f_{o} = \frac{\sin k_{\nu}}{k_{\nu}}$$

$$f_{1} = i \left(\frac{\sin k_{\nu}}{k_{\nu}^{2}} - \frac{\cos k_{\nu}}{k_{\nu}} \right)$$

$$f_{2} = \frac{3 \cos k_{\nu}}{k_{\nu}^{2}} - \left(\frac{6 - 3k_{\nu}^{2}}{2k_{\nu}^{3}} + \frac{1}{2k_{\nu}} \right) \sin k_{\nu}$$

$$f_{3} = i \left(\left(\frac{30 - 5k_{\nu}^{2}}{2k_{\nu}^{3}} + \frac{3}{2k_{\nu}} \right) \cos k_{\nu} - \left(\frac{30 - 15k_{\nu}^{2}}{2k_{\nu}^{4}} + \frac{3}{2k_{\nu}^{2}} \right) \sin k_{\nu} \right)$$

$$f_{4} = \left(\frac{35(k_{\nu}^{2} - 6)}{2k_{\nu}^{4}} - \frac{15}{2k_{\nu}^{2}} \right) \cos k_{\nu} + \left(\frac{35(k_{\nu}^{4} - 12k_{\nu}^{2} + 24)}{8k_{\nu}^{5}} + \frac{30(2 - k_{\nu}^{2})}{8k_{\nu}^{3}} + \frac{3}{8k_{\nu}} \right) \sin k_{\nu}$$

$$f_{5} = i \left(\frac{-2k_{\nu}^{4} + 210k_{\nu}^{2} - 1890}{k_{\nu}^{5}} \cos k_{\nu} + \frac{30k_{\nu}^{4} - 840k_{\nu}^{2} + 1890}{k_{\nu}^{6}} \sin k_{\nu} \right)$$

$$f_{6} = \left(\frac{42k_{\nu}^{4} - 2520k_{\nu}^{2} + 20790)}{k_{\nu}^{6}} \right) \cos k_{\nu} + \left(\frac{2k_{\nu}^{6} - 420k_{\nu}^{4} + 9450k_{\nu}^{2} - 20790)}{k_{\nu}^{7}} \right) \sin k_{\nu}$$

$$(8.75)$$

We now see that there is a natural extension of this idea to multiple baselines, adding two new structure parameters per baseline, so that in general N baselines yields 2N + 1 terms of the Fourier–Legendre series. Returning to the N = 2 case, CT inversion then takes the form of a matrix equation based on the use of equation (8.74) for two baselines 'x' and 'y', as shown in equation (8.76):

$$\begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & -if_1^x & 0 & -if_3^x & 0 \\ 0 & 0 & f_2^x & 0 & f_4^x \\ 0 & -if_1^y & 0 & -if_3^y & 0 \\ 0 & 0 & f_2^y & 0 & f_4^y \end{bmatrix} . \begin{bmatrix} a_{00} \\ a_{10} \\ a_{20} \\ a_{30} \\ a_{40} \end{bmatrix} = \begin{bmatrix} 1 \\ \operatorname{Im}\left(\tilde{\gamma}_k^x\right) \\ \operatorname{Re}\left(\tilde{\gamma}_k^y\right) - f_0^y \\ \operatorname{Re}\left(\tilde{\gamma}_k^y\right) - f_0^y \end{bmatrix} \Rightarrow \hat{\underline{a}} = [L]^{-1}\underline{\underline{b}}$$

(8.76)

Note that the real matrix [F] is now 5 × 5 (for N baselines it is $(2N + 1) \times (2N + 1)$), and is no longer diagonal in structure. From the estimated vector of Legendre coefficients we can determine the shape of the corresponding structure function up to fourth order, as shown in equation (8.77):

$$\hat{f}(z') = 1 + \hat{a}_{10}P_1(z') + \hat{a}_{20}P_2(z') + \hat{a}_{30}P_4(z') + \hat{a}_{40}P_4(z') - 1 \le z' \le 1$$
(8.77)

Extending this to N = 3 (to three baselines) leads to the following model for coherence, where again the new functions f_5 and f_6 are given in equation (8.75).

$$\tilde{\gamma} e^{-ik_{\gamma}} e^{-i\phi_o} = \tilde{\gamma}_k$$

$$= f_0 + a_{10}f_1 + a_{20}f_2 + a_{30}f_3 + a_{40}f_4 + a_{50}f_5 + a_{60}f_6 \quad (8.78)$$

When applied across the three baselines 'x', 'y' and 'z', this leads to a corresponding 7×7 matrix inversion for coherence tomography, as shown in equation (8.79). Note again that [*F*] is a real non-diagonal matrix with elements a function of k_v .

$$\begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -if_1^x & 0 & -if_3^x & 0 & -if_5^x & 0 \\ 0 & 0 & f_2^x & 0 & f_4^x & 0 & f_6^x \\ 0 & -if_1^y & 0 & -if_3^y & 0 & -if_5^y & 0 \\ 0 & 0 & f_2^y & 0 & f_4^y & 0 & f_6^y \\ 0 & -if_1^z & 0 & -if_3^z & 0 & -if_5^z & 0 \\ 0 & 0 & f_2^x & 0 & f_4^x & 0 & f_6^x \end{bmatrix} \begin{bmatrix} a_{00} \\ a_{10} \\ a_{20} \\ a_{30} \\ a_{40} \\ a_{50} \\ a_{60} \end{bmatrix} = \begin{bmatrix} 1 \\ \operatorname{Im}\left(\tilde{\gamma}_k^x\right) \\ \operatorname{Re}\left(\tilde{\gamma}_k^x\right) - f_0^x \\ \operatorname{Im}\left(\tilde{\gamma}_k^y\right) \\ \operatorname{Re}\left(\tilde{\gamma}_k^x\right) - f_0^y \\ \operatorname{Im}\left(\tilde{\gamma}_k^z\right) \\ \operatorname{Re}\left(\tilde{\gamma}_k^z\right) - f_0^z \end{bmatrix}$$
(8.79)

This then permits an even higher-resolution reconstruction of the structure function, as shown in equation (8.80):

$$\hat{f}(z') = 1 + \hat{a}_{10}P_1(z') + \hat{a}_{20}P_2(z') + \hat{a}_{30}P_4(z') + \hat{a}_{40}P_4(z') + \hat{a}_{50}P_5(z') + \hat{a}_{60}P_6(z')$$
(8.80)

Figure 8.39 summarizes the differences between single, dual and triple baseline reconstructions by plotting the polynomials employed in the corresponding reconstructions. We can clearly see the improvement in resolution with increasing baseline. However, while formulation of CT in this way is straightforward, note from Figure 5.15 that the functions f_i tend to zero with increasing order and hence anticipate problems with the conditioning of the inversion. This will provide a practical limit to the achievable resolution, as eventually we will demand impossible limits on the control of error in coherence estimation for the vector <u>b</u>. In order to assess this, we now turn to quantify the condition number of multi-baseline CT.

In general we can analyse the conditioning of multi-baseline CT using a singular value decomposition of the matrix [L]. This allows us to represent the inversion in terms of a $(2N + 1) \times (2N + 1)$ diagonal matrix [Σ], just as we did for the single-baseline case in equation (8.68), but now with different



Fig. 8.39 Legendre functions for single, dual, and triple baseline inversions

orthogonal frames [U] and [V] for the vectors \underline{a} and \underline{b} . Formally we can write the matrix [L] in the form shown in equation (8.81) (see Appendix 1):

$$[L] = [U] \cdot [\Sigma] \cdot [V]^{*T} \quad s_1 \ge s_2 \ge \dots \ge s_{2N+1}$$
(8.81)

where the (2N + 1) real parameters s_i are the singular values. The formal solution to the CT estimation problem can be written in terms of these matrix components, as shown in equation (8.82):

$$\hat{\underline{a}} = [V] \cdot [\Sigma]^{-1} [U]^{*T} \underline{b},$$

$$[\Sigma]^{-1} = \begin{bmatrix} s_1^{-1} & 0 & 0 & \dots & 0 \\ 0 & s_2^{-1} & 0 & \dots & 0 \\ 0 & 0 & s_3^{-1} & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & s_{2N+1}^{-1} \end{bmatrix}$$
(8.82)

The condition number of the inversion is then defined as the ratio of maximum to minimum singular values, as shown in equation (8.83):

$$CN = \frac{s_1}{s_{2N+1}}$$
(8.83)

As an example, we show in Figure 8.40 the variation of CN on a dB scale for baseline pairs $k_{v1}.k_{v2}$ in dual-baseline CT. Note that along the diagonal the CN goes to infinity, since the rows of [*F*] no longer provide independent information about the structure. We see that in useful portions of the k_v space (around 1) the CN is very high, of the order of 1000 or more.



Fig. 8.40 Variation of condition number for dual-baseline inversions with baselines kvx and kvy

Fig. 8.41 Variation of condition number for singular value filtered dual-baseline inversions with baselines k_{vx} and k_{vy}

One way to deal with this high condition number is to filter the [L] matrix, which can be achieved by removing the smallest singular value to reconstruct a profile with a matrix $[\Sigma_f]$, as shown in equation (8.84). Here we obtain a matrix with a lower condition number, given by s_1/s_{2N} . Note that in this case we lose some resolution (given by one pair of singular vectors), but still gain resolution over the reduced baseline case. Figure 8.41 shows the condition number of this filtered matrix. Here we see an order of magnitude improvement in conditioning, with condition numbers around 100 for the useful part of the $k_{\rm v}$ range.

$$\hat{\underline{a}}_{f} = [V] \cdot [\Sigma_{f}]^{-1} [U]^{*T} \underline{b},$$

$$[\Sigma_{f}]^{-1} = \begin{bmatrix} s_{1}^{-1} & 0 & \dots & \dots & 0\\ 0 & s_{2}^{-1} & \dots & \dots & 0\\ \vdots & \vdots & \ddots & \vdots & \vdots\\ 0 & 0 & 0 & s_{2N}^{-1} & 0\\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$
(8.84)

We will provide an example of dual and triple baseline CT processing of anechoic chamber data in Chapter 9.

9

Applications of polarimetry and interferometry

In this chapter we turn to consider some applications of polarimetry and polarimetric interferometry in remote sensing. A comprehensive survey would be impossible, and so instead we select a few representative examples taken from different areas. We do this firstly to reinforce the theoretical ideas introduced in earlier chapters, but also to present an idea of the wide range of topics in which these concepts can be applied. We start with a general introduction to synthetic aperture radar (SAR) (Curlander, 1991; Mensa, 1991), as it is with radar imaging that most applications currently occur. In particular we outline a hierarchy of polarimetric modes in radar imaging, starting with single-channel SAR and then interferometric SAR, or InSAR (Bamler, 1998), before developing into both compact and quad polarimetric, or POLSAR (Lee, 2008; Kong, 1990; Mott, 2007; Ulaby, 1990), and finally to imaging polarimetric interferometry. or POLInSAR (Cloude, 1998, 2001b; Krieger, 2005).

We then turn to consider several application themes, starting with bare surface scattering and then considering the effects of vegetation cover, first through agriculture or short vegetation and then considering the important case of forestry. We finally turn to consider applications centred around the study of isolated point scatterers, such as occur in urban areas and in ship detection and monitoring. In this way we cover a broad range of topics that illustrate many of the concepts introduced in earlier chapters.

9.1 Radar imaging

We begin by considering the basic principles of radar imaging. More details can be found in the specialist monographs by Curlander (1991), Mensa (1991), and Franceschetti (1999). Consider a static transmitter/receiver configuration as shown schematically in Figure 9.1. When we employ a transmitter and receiver separated by bistatic angle Φ and operating at a single wavelength λ , then scattering by the environment around the transmitter leads to a total received signal in amplitude and phase, represented by a complex number. This complex number in fact represents the amplitude of a Fourier component located at point P in a wavenumber space, as shown on the right-hand side of Figure 9.1. The polar coordinates of the point P in this space are then defined by the geometry of the transmitter and receiver configuration and the propagation phase delay between the two, defined as $\exp i\beta(\underline{r}_1 + \underline{r}_2)$), leading to the triangular construction shown in Figure 9.1. Clearly such a single static configuration does not lead to an image of the environment. The signal obtained in the receiver is the coherent summation from many points in the scene, depending on many factors



Fig. 9.1 Wave-space geometry for a single transmitter/receiver combination

Fig. 9.2 Synthetic aperture geometry

including the beamwidth of the transmitter and receiver antennas. To obtain an image requires diversity over one or more of the three parameters λ , θ and Φ in order to fill a sector of wave space. When such a sector has been filled, then an inverse two-dimensional Fourier transform can be used to reconstruct an image of the environment, as we now formally demonstrate.

By far the most common radar configuration is to employ a monostatic sensor ($\Phi = 0$) working in backscatter, with a finite bandwidth *W* representing wavelength diversity, and then linear motion of the radar system to generate θ diversity. The latter generates a finite line segment along the radar flight trajectory, as shown by AA' in Figure 9.2. This segment can be considered a *synthetic* antenna aperture (which is much larger than the actual physical antenna aperture)—hence the term synthetic aperture radar, or SAR.

A two-dimensional image of the environment in the z-x plane can then be obtained using an $\omega - \beta$ SAR processor as follows. (Note that in many texts this is called an $\omega - k$ processor, where k is the symbol for wavenumber. However, here we use <u>k</u> for scattering vector, and so to avoid confusion we refer to the $\omega - \beta$ processor.) In Figure 9.2, T is a general reflecting point in the scene, O is the (monostatic) radar observation point and AA' the linear 'aperture' of the radar flight path. If we denote the wave field caused by an apparent source at T as d(x, z, t) then we know that d everywhere obeys the following wave equation (Gazdag, 1984; Cafforio, 1991):

$$\frac{\partial^2 d}{\partial x^2} + \frac{\partial^2 d}{\partial z^2} = \frac{1}{v^2} \frac{\partial^2 d}{\partial t^2}$$
(9.1)

If we now take the Fourier transform of this equation with respect to time and to *x*, we obtain the following (ordinary) differential equation for the transform quantity $D(\beta_x, \omega, z)$:

$$\beta_x^2 D - \frac{d^2 D}{dz^2} = \frac{\omega^2}{v^2} D \Rightarrow \frac{d^2 D}{dz^2} = -\frac{\omega^2}{v^2} (1 - \frac{\beta_x^2 v^2}{\omega^2}) D$$
(9.2)

This 'ODE' can then be factored in terms of upward and downward $(\pm z)$ propagating waves. The latter we can then use to 'migrate' the field from the line AA' back to the source line at $z = z_0$. This will render the wave field sensed along AA' as an 'image' of the apparent 'sources' along $z = z_0$; that is, it will focus the radar image. We shall see that the larger the aperture AA', the higher the resolution in this image. For the downward propagating waves we have the following factorization:

$$\frac{dD}{dz} = f.D = i\frac{\omega}{v}\sqrt{1 - \left(\frac{\beta_x v}{\omega}\right)^2} D = i\frac{\omega}{v}\sqrt{1 - \eta^2}D$$
(9.3)

This equation has a simple plane wave solution. Therefore, if we know D across a surface AA' then we can propagate or 'migrate' the data to any other z value, as shown in equation (9.4):

$$D(\beta_x, \omega, z + \Delta z) = D(\beta_x, \omega, z)e^{i\frac{\omega}{\nu}\sqrt{1 - \eta^2 \Delta z}}$$
(9.4)

Finally, we can obtain the image of the sources by an inverse Fourier transform (FT) with respect to β_x and a summation w.r.t. ω as follows:

$$d(x,t=0,z) = \sum_{k_x} \sum_{\omega} D(\beta_x,\omega,z) e^{i\beta_x x}$$
(9.5)

If the velocity is constant (v = c/2 to account for the two-way propagation) then this summation can be performed very efficiently by using a two-dimensional Fourier transform as follows. We first propagate the data from AA' to $z = z_o$ in one single step by phase rotation, as shown in equation (9.6):

$$D(\beta_x, \omega, z_o) = D(\beta_x, \omega, z = 0)e^{i\beta_z z_o} \quad \beta_z = \frac{\omega}{v}\sqrt{1 - \eta^2}$$
(9.6)

We then perform the summation and inverse transform to obtain the following integral:

$$d(x,t=0,z_o) = \iint D(\beta_x,\omega,z=0).e^{i(\beta_z(\omega)z+\beta_x x)}d\beta_x d\omega \qquad (9.7)$$

This is almost a two-dimensional FT operation, and we can complete the process by a change of variable $\omega = \frac{c}{2}\sqrt{\beta_x^2 + \beta_z^2}$ and integration with respect to β_z instead of ω to obtain the following Fourier transform relationship between the measured spectral function *D* and original source distribution:

$$d(x,t=0,z_o) = \iint \frac{c\beta_z}{\sqrt{\beta_x^2 + \beta_z^2}} D(\beta_x,\omega,z=0).e^{i(\beta_z z + \beta_x x)} d\beta_x d\beta_z \quad (9.8)$$

This approach of wave migration to generate a SAR image is called the wavenumber processor. It is only one of several approaches to SAR processing (Bamler, 1992, 1998; Curlander, 1991), but for our purposes provides a direct link to the wave equation and propagating polarised EM waves. As shown above, there are three major steps involved in the ω - β processor.

- Collect raw signal data d(x, t, z = 0) and perform a two-dimensional Fourier transform (FT) to obtain $D(\beta_x, \omega, z = 0)$. In practice this stage can be very efficiently implemented by using coherent IQ sampling of the signal and a digital signal processor.
- Evaluate the complex function *D* over a regular grid in β_z , β_x (called Stolt interpolation).
- Multiply by the Jacobian and inverse two-dimensional FT to obtain a two-dimensional image.

Note that only two parameters are important for correct focusing of the image:

- The platform velocity $v_p(\beta_x \text{ depends on } v_p)$.
- The parameter z_0 —the distance to the front of the range gate used.

Hence both of these need to be known accurately in order to focus the image correctly. This basic $\omega - \beta$ SAR processor (we have ignored, for example, important practical issues such as motion compensation, by assuming that the sensor moves in a perfect straight line) is summarized geometrically in Figure 9.3. The resulting image is complex, as at each pixel we obtain an estimate of the scattering in both amplitude and phase from that point. The resolution we obtain depends on the angular and radial extent of the measured sector in wave space. It is common to process a narrow sector $\Delta \theta$ centred around $\theta = 0^{\circ}$ (by pointing the real antenna axis at right angles to the flight vector). In this case the resolution in z (range) and cross-range or azimuth (x) can be simply related to two system parameters: the transmitter bandwidth W, and the real antenna dimension in the along-track or x direction. To see this we use the relationship between Fourier transform variables and estimates of the bandwidth in wave space in both the z and x directions, as shown in Figure 9.4. Here we see that the resolution in the range direction depends only on the bandwidth W of the transmitted signal, while the cross-range resolution



Fig. 9.3 Schematic geometry and key steps in SAR processing



Fig. 9.4 Resolution in SAR imaging



Fig. 9.5 Approximate expression for antenna beamwidth

depends on the angular width of the sector. At first sight this may seem to depend on several system variables, but there is a simple relationship between the maximum angular width (and hence best resolution) and real antenna size, as shown schematically in Figure 9.5. The beamwidth (in radians) of the real antenna is approximately equal to the size of the aperture in wavelengths. When we substitute this result into Figure 9.4 we obtain the well-known result, from SAR theory, that the cross-range resolution is given by half the real antenna aperture size:

$$\Delta \theta \approx \frac{\lambda}{d_r} \Rightarrow \Delta x \approx \frac{d_r}{2} \tag{9.9}$$

This result has important implications for the exploitation of polarisation effects in radar imaging, as we now show.

9.1.1 PRF, antenna size and Doppler bandwidth

From equation (9.9) we see that the smaller the antenna, the wider its beamwidth, the wider the measured sector in wavespace, and hence the better the resolution. However, since SAR involves a sampled measurement system on a moving platform, we must be careful that the sampling of phase across the aperture is performed fast enough so as to avoid any sampling errors due to aliasing. The pulses are transmitted at a rate called the PRF or pulse repetition frequency and, to avoid sampling errors, this PRF must be greater than or equal to the maximum rate of change of phase. The time rate of change of phase across the aperture is just the Doppler frequency of the received signal due to the relative motion between radar and sample point. Doppler shift is zero when the velocity vector is perpendicular to the line of sight vector to the point—at the centre of the antenna pattern in side-looking geometry. In general, the Doppler shift of the signal from a point with angular position θ inside the beam is proportional to $\sin\theta$, as shown on the left-hand side of equation (9.10):

$$f_d = \frac{4\nu\sin\theta}{\lambda} \Rightarrow f_{d\max} \approx \frac{4\nu}{\lambda} \cdot \frac{\Delta\theta}{2} = \frac{2\nu}{\lambda} \cdot \frac{\lambda}{d_r} = \frac{2\nu}{d_r}$$
(9.10)

The maximum Doppler shift therefore occurs at the outer edges of the real beam (positive for approaching points, and negative for receding). By again using our approximation for the beamwidth of the antenna in terms of the real aperture size (Figure 9.5), we can obtain a simple expression for the maximum Doppler

shift as a function of the ratio of platform speed to real aperture size, as shown on the right-hand side of equation (9.10). One important constraint required for undistorted SAR imaging is therefore that the PRF be greater than or equal to this maximum shift. For small antennas on fast-moving platforms this can require a very high PRF. However, there are two consequences of operating at high PRF. The first is the requirement for a transmitter with higher mean power (given by the peak power times the duty cycle of the radar or τ *PRF, where τ is the pulse width), which may be expensive or difficult to obtain at the desired operating frequency. The second issue, however, is more important for imaging, in that the PRF also impacts on the range extent of the image or range swath in pulsed systems. The problem here arises from range ambiguities. If the PRF is too high and the range variation across the image too large, then there can be an ambiguity as to which transmitter pulse any particular received pulse actually belongs. A quantitative analysis (Curlander, 1991) shows that the PRF must be bounded by the following inequality in order to avoid such range ambiguities:

$$PRF \le \frac{c}{2W_s} \tag{9.11}$$

where W_s is the width of the image in the range direction (the range swath size). This tends to demand low PRF for wide image coverage of the system, which, as we have seen, is in direct contrast with the requirements for high resolution in the cross-range direction. The compromise between such conflicting PRF requirements is one of the central engineering steps in imaging radar design. Polarisation switching places further constraints on this relationship, as we show in Section 9.3; but first we turn to consider radar interferometry.

9.2 Imaging interferometry: InSAR

The above ideas can be extended to imaging interferometry by combining two SAR images generated by linear trajectories separated by a baseline vector b. Figure 9.6 shows a schematic representation of this process. The two tracks will in general fill different sectors of wavenumber space, shown as θ_1 and θ_2 in the figure. By applying a two-dimensional inverse Fourier transform (IFT) to the separate images we obtain two complex images. However, for successful interferometry we require good coherence between the two images, and so the same regions should be processed to generate the two images. In general, the two regions will overlap only over part of their wave space coverage, and this will reduce the resolution. This is called common-band filtering, and we see that in the imaging context it is a two-dimensional process. In the azimuth or x-direction we require there to be an angular sector of the same width and with the same mean. This implies that the same squint of the real antenna be used. A common approach is to employ zero squint; that is, to process to zero Doppler in both images. This then maintains coherence, and maximizes overlap and hence resolution. In the z or range direction the radars should have the same carrier frequency and bandwidth; but we also note that in this direction range, spectral filtering will be required to shift the pulse bandwidth of the second track so as to remove baseline decorrelation according to the discussion in Section 5.1.1.1.



By co-registering images so that the point P has exactly overlapping pixels in the two image spaces, we obtain an image with resolution given by the SAR process, at each pixel of which we can generate a phase difference. In this way we generate a high-resolution interferometer whereby we can track spatial changes in interferometric phase and coherence across a scene. Under assumptions of stationarity and ergodicity we can then estimate the mean interferometric phase and coherence using a rectangular window in image space, centred on the pixel of interest, as shown in Figure 9.7 (Touzi, 1999).

$$\tilde{\gamma} = \frac{\sum_{i=1}^{MN} s_{1i} s_{2i}^*}{\sqrt{\sum_{i=1}^{MN} s_{1i} s_{1i}^*} \sqrt{\sum_{i=1}^{MN} s_{2i} s_{2i}^*}} \begin{cases} 0 \le |\tilde{\gamma}| \le 1\\ 0 \le \arg(\tilde{\gamma}) < 2\pi \end{cases}$$
(9.12)

If the window size is $M \times N$ pixels, we have MN samples available for coherence estimation. Clearly, by using large windows we can secure more accurate estimates of coherence, but at the same time are reducing the effective resolution of the image. This idea of multiple channel imaging and combining channels coherently is also employed in polarimetric SAR, as we consider in the next section.

9.3 Polarimetric synthetic aperture radar (POLSAR)

The extension of the SAR concept to the polarimetric case is in principle straightforward. In place of a single complex number at each location in wave space, we require a set of four complex numbers representing the scattering vector at that wave space coordinate. This is shown schematically on the left-hand side of Figure 9.8. Repeating the SAR imaging process (the $\omega - \beta$ algorithm of equation 9.8) for each of the four channels separately leads to four images—one for each component of the scattering matrix, as shown on the right of Figure 9.8. We can then take linear combinations of the (complex) elements using the wvector concept to form an image of scattering mechanism w. We can also study local variations in depolarisation by estimating the coherency matrix from a weighted sum about the pixel of interest, as shown schematically in Figure 9.9. Note that this assumes stationarity and ergodicity in that the spatial locality of the pixel is assumed to consist of random samples from an underlying stochastic process with the same coherency matrix. Under this assumption we can then estimate the coherency (or Mueller) matrix locally and apply eigenvalue decomposition or any of the other processing techniques discussed in Chapter 4. There are two important points to note:

1. The coherency matrix obtained in this way is only ever an estimate, usually obtained from a relatively small number of samples (depending on





Fig. 9.8 Wave-space interpretation of POL-SAR imaging





Fig. 9.9 Pixel averaging of POLSAR data
window size $M \times N$). Hence it contains estimation errors due to speckle fluctuations (Lee, 1994a), and these must be accounted for using, for example, the multivariate Wishart distribution in any quantitative assessment of the elements of [*T*] (see Appendix 3) (Conradson, 2003; Schou, 2003; Ferro-Famil, 2008).

2. The weights for each pixel inside the window need not be unity (which corresponds to the standard so called 'boxcar' filter). One reason for varying the elements is that such a window degrades the effective resolution of the image. The convolution of the rectangular window shape with the image is equivalent to multiplication of the corresponding Fourier spectra and hence to a low-pass filtering of the image with a SINC reference spectrum with a width inversely proportional to window size. A better approach is often to adaptively estimate the weights over the image, using an estimate of *local* statistics (around the pixel of interest). The most popular form of such locally adaptive filtering in radar imaging is the Lee filter (Lee, 1999, 2008). This filter forms an estimate of coherency matrix from local samples according to the following weighted contributions:

$$\hat{T} = \langle T \rangle + f(\underline{k}_i \underline{k}_i^{*T} - \langle T \rangle)$$
(9.13)

where f is to be determined from the local statistics. In homogeneous areas (areas with fully developed speckle), f = 0 and the average matrix is taken as the estimate. On the other hand, for inhomogeneous areas (isolated point scatterers, for example) f = 1, and the estimate is obtained using only the central pixel itself, so preserving spatial resolution. Note that in order to preserve the correct polarimetric information in the coherency matrix, the same f should be used on all elements of the T matrix. Details of the expression for f in terms of local statistics can be found in Appendix 3.

As a popular extension of the Lee filter, the window shape itself is modified to account for edges at 0° and 45° to the image boundaries. This leads to a family of eight Lee filters for each pixel, with the best matched to the local scene being selected (Lee, 2008). This additional complexity is used in an attempt to further improve the balance between preservation of spatial resolution in heterogeneous parts of the image (such as at edges, and for point scatterers) while maintaining good radiometric resolution (reducing estimation bias of coherency matrices) in homogeneous regions.

All of this forms a natural extension of single-channel imaging, but we have so far assumed that all four elements of the scattering matrix can be measured simultaneously for all points in wave space. In practice this is not possible, and the coding schemes employed have important implications for the PRF of the imaging radar, as we now consider.

9.3.1 Pulse switching requirements for POLSAR imaging

Measurement of the four complex matrix elements of [S] requires transmission of two orthogonal polarisations x and y, represented by end points of a diameter of the Poincaré sphere. In principle, x and y can be any orthogonal pair, but the most common selections are horizontal and vertical linear (H and V) or

$$H \qquad \begin{bmatrix} S \end{bmatrix} = \begin{bmatrix} S_{HH} & S_{VH} \\ S_{HV} & S_{VV} \end{bmatrix} \qquad V \qquad \begin{bmatrix} S \end{bmatrix} = \begin{bmatrix} S_{HH} & S_{VH} \\ S_{HV} & S_{VV} \end{bmatrix} \qquad H \qquad \begin{bmatrix} S \end{bmatrix} = \begin{bmatrix} S_{HH} & S_{VH} \\ S_{HV} & S_{VV} \end{bmatrix} \qquad V \qquad Fig. 9.10 \text{ Pulse swite}$$

Fig. 9.10 Pulse switching in quadpol systems

left and right circular (L and R). In order to measure the first column of [S] we illuminate the scatterer with x polarisation and measure, simultaneously in amplitude and phase, the scattered field components in the orthogonal x and y channels. Simultaneous dual reception can be achieved using a two-channel receiver preceded by an orthogonal polarisation splitter, although, as we shall see, this complicates the calibration, as multiple channels need to be balanced in both amplitude and phase (Freeman, 1992).

The second column of [S] can be similarly measured by illuminating the scatterer with orthogonal y polarisation and again measuring coherently the x and y components of the scattered radiation. In this way, all four complex matrix elements are obtained. Ideally, as SAR involves a moving platform we should transmit x and y polarisations simultaneously. This could in principle be achieved using suitable orthogonal coding. However, by far the most common method is to employ a single carrier frequency and time multiplex the two orthogonal states on a pulse-by-pulse basis, as shown schematically in Figure 9.10. Here we first transmit a horizontal or H polarised pulse, and receive in the co- and crosspolarised channels the first column of [S], as shown. The next pulse is then transmitted with V polarisation, and we measure the second column of [S]. In this switching scheme there is an inherent time delay of one PRI (pulse repetition interval) between the first and second column, and so the bandwidth of the transmitter switch needs to be much faster than any decorrelation time of the scattering process, so as to maintain coherence between the columns of [S]. Bandwidths in the kHz region are typical for imaging radar applications.

We see, however, that this interleaved switching arrangement also interferes with the sampling requirements for SAR processing. There are two main options. The first is to keep the same PRF (and hence mean power) as for a single polarisation system. However, this means that the effective PRF for each column measurement is halved, as H is transmitted only on every second pulse in Figure 9.10. This in turn means that the azimuth resolution is also halved, and a larger antenna (with twice the size in the azimuth direction) is required to avoid Doppler aliasing. Both of these are unattractive options for imaging radar systems.

In the second scenario we can instead double the PRF of the system so as to maintain the column sampling at the same rate as before, and so maintain azimuth resolution and keep the same antenna size. However, in this case the mean power of the transmitter is doubled (unless the pulse length is also halved to maintain the same duty cycle, which then has further implications for the range resolution, which may have to be halved). This option also leads to a halving of the range swath and hence a smaller image, due to the possibility of range ambiguities between columns of [S] (especially in the crosspolarised channels). In equation (9.11), therefore, we need to use the full PRF in determining the range swath, leading to reduction by a factor of 2. All these considerations are worsened by the fact that systems always have finite isolation between orthogonal channels; that is, there will inevitably be some y component radiated when



Fig. 9.11 Calibration diagram for transmitter and receiver distortions in radar polarimetry

x is selected, and *vice versa*. Methods for dealing with such practical issues via calibration are dealt with in the next section.

9.3.2 Polarimetric calibration

Practical devices and systems are never perfect, and there will inevitably be some corruption of the measured scattering matrix elements by system imperfections due, for example, to undesired cross-talk between channels, and amplitude or phase imbalance in transmitter and receiver systems (van Zyl, 1990; Freeman, 1992; Sarabandi, 1992b; Quegan, 1994). To quantify such distortions we can employ a cascade of matrices in a composite product, as shown in equation (9.14). Figure 9.11 shows how this distortion chain originates. First the 'ideal' orthogonal states x and y are passed through the transmitter chain (including antenna), which incurs some distortions via the channel imbalance terms t_{11} and t_{22} , as well as undesired cross-talk via t_{12} and t_{21} .

This transmitted wave then interacts with the scatterer, and the desired changes in amplitude and phase caused by scattering are imprinted on the signal. On return to the receiver there is yet another series of distortions and the addition of thermal noise in the receiver before the observed matrix elements O_{ij} are obtained. We can formulate an expression for all these processes based on matrix multiplication, as shown in equation (9.14). This is a standard model widely used for calibration of polarimetric radar systems (Papathanassiou, 1998a; Kimura, 2004).

$$[O] = \begin{bmatrix} r_{11} & r_{12} \\ r_{21} & r_{22} \end{bmatrix} \begin{bmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{bmatrix} \begin{bmatrix} t_{11} & t_{12} \\ t_{21} & t_{22} \end{bmatrix} + \begin{bmatrix} n_{11} & n_{12} \\ n_{21} & n_{22} \end{bmatrix}$$
$$= [R] \ [S] \ [T] + [N] \tag{9.14}$$

There are two strategies for dealing with these distortions. In design, every effort can be made to reduce cross-talk by good antenna design and careful system layout. As a second strategy we can employ the process of calibration to estimate the distortion matrices [R] and [T], and remove them by matrix inversions, so that in the absence of noise, for example, we can obtain an estimate of the true scattering matrix, as shown in equation (9.15):

$$[S] = [R]^{-1} [O] [T]^{-1}$$
(9.15)

There are various methods available for the estimation of the elements of [R] and [T], involving a combination of internal and external calibration techniques. Internal methods involve monitoring of signals by test channels inside the radar to estimate imbalances, while external methods (which have the advantage that they include the full system, including antenna and propagation effects) involve measuring signals from external active and passive reflectors, which send signals back through the radar system from an object with *known* polarimetric behaviour. By arranging for a set of four such reflectors with orthogonal scattering vectors (see Section 4.1.4), a set of sixteen equations in the sixteen unknowns of [R] and [T] can be obtained. In practice, simpler deployments are favoured, often of only one or two types of reflector with additional constraints (such as symmetry assumptions in the scattering from random media; see Section 2.4.2.1) used to solve the remaining calibration equations. To see how these arise we now turn to a vector formulation of the system calibration equations.

9.3.3 Scattering vector formulation of polarimetric calibration

One important application of the scattering vector formulation is in the treatment of polarimetric system calibration. We showed in equation (9.14) how the distortions due to system imperfections can be represented as a triple matrix product. Ignoring noise, and using the expansion of such a product into a single matrix equation, we obtain the scattering vector distortion matrix [Z], as shown in equation (9.16):

$$\begin{bmatrix} O \end{bmatrix} = \begin{bmatrix} R \end{bmatrix} \begin{bmatrix} S \end{bmatrix} \begin{bmatrix} T \end{bmatrix} \Rightarrow$$

$$\underline{k}_{obs} = \begin{bmatrix} O_{HH} \\ O_{HV} \\ O_{VH} \\ O_{VV} \end{bmatrix} = \begin{bmatrix} Z \end{bmatrix} \underline{k}_{s} = \begin{bmatrix} r_{11}t_{11} & r_{11}t_{21} & r_{12}t_{11} & r_{12}t_{21} \\ r_{11}t_{12} & r_{11}t_{22} & r_{12}t_{12} & r_{12}t_{22} \\ r_{21}t_{11} & r_{21}t_{21} & r_{22}t_{11} & r_{22}t_{21} \\ r_{21}t_{12} & r_{21}t_{22} & r_{22}t_{12} & r_{21}t_{22} \end{bmatrix} \cdot \begin{bmatrix} S_{HH} \\ S_{HV} \\ S_{VH} \\ S_{VV} \end{bmatrix}$$

$$\Rightarrow \underline{k}_{s} = \begin{bmatrix} Z \end{bmatrix}^{-1} \underline{k}_{obs}$$
(9.16)

The two key features of this formulation are the presence of quadratic products of the distortion matrices appearing in [*Z*], and the simple mathematical form of the correction process. If we can estimate the elements of [*R*] and [*T*], then their distortions can be offset by a single matrix inversion as shown. There are two important special forms of this calibration matrix to consider, both of which stem from the important case of reciprocal backscatter when $S_{HV} = S_{VH}$. In this case the matrix [*Z*] is no longer square and has dimension 4×3 , as shown in equation (9.17):

$$\underline{k}_{obs} = \begin{bmatrix} O_{HH} \\ O_{HV} \\ O_{VH} \\ O_{VV} \end{bmatrix} = [Z] \underline{k}_{s} = \begin{bmatrix} r_{11}t_{11} & r_{11}t_{21} + r_{12}t_{11} & r_{12}t_{21} \\ r_{11}t_{12} & r_{11}t_{22} + r_{12}t_{12} & r_{12}t_{22} \\ r_{21}t_{11} & r_{21}t_{21} + r_{22}t_{11} & r_{22}t_{21} \\ r_{21}t_{12} & r_{21}t_{22} + r_{22}t_{12} & r_{22}t_{22} \end{bmatrix} \cdot \begin{bmatrix} S_{HH} \\ S_{HV} \\ S_{VV} \end{bmatrix}$$
$$\Rightarrow \underline{k}_{s} = (Z^{*T}Z)^{-1} Z^{*T} \underline{k}_{obs}$$
(9.17)

Note that the *observed* scattering vector \underline{k}_{obs} violates reciprocity, though this is due entirely to the effect of system distortions. This observation can be used,

for example, to test the quality of system calibration. If the Pauli channel image $O_{HV} - O_{VH}$ is formed, it should, for properly calibrated backscatter data, behave like noise (have zero coherence, and so on). If it contains structure, then the calibration is not perfect. Note that calibration of the data does not now involve matrix inversion directly, but instead a pseudo-inverse based on a least squares solution can be employed, as shown in equation (9.17). This arises because we are using reciprocity to reduce the number of unknowns below the number of observations, and hence have an overdetermined system of equations.

In practice, a further simplification can often be made by assuming that the cross-talk terms (the off-diagonal elements of [R] and [T]) are small compared to the copolar distortion terms. In this case we can set to zero elements of the [Z] matrix that involve products of the small crosspolar terms. In this 'small-coupling' assumption the [Z] matrix takes on the simplified form shown in equation (9.18):

$$[Z] = \begin{bmatrix} r_{11}t_{11} & r_{11}t_{21} + r_{12}t_{11} & 0\\ r_{11}t_{12} & r_{11}t_{22} & r_{12}t_{22}\\ r_{21}t_{11} & r_{22}t_{11} & r_{22}t_{21}\\ 0 & r_{21}t_{22} + r_{22}t_{12} & r_{22}t_{22} \end{bmatrix}$$
(9.18)

A special case of this matrix occurs in the limit of zero cross-talk. If the design isolation of the system is very good (typically better than -30 dB), then we can set all off-diagonal terms of [*R*] and [*T*] to zero and establish a simplified calibration matrix as shown in equation (9.19):

$$[Z] = r_{11}t_{11} \begin{bmatrix} 1 & 0 & 0 \\ 0 & t_{22}/t_{11} & 0 \\ 0 & r_{22}/r_{11} & 0 \\ 0 & 0 & r_{22}t_{22}/r_{11}t_{11} \end{bmatrix}$$
(9.19)

Note that this matrix still causes lack of reciprocity in the observed vector due to differences in the receiver and transmitter copolar distortion channels.

To illustrate how this matrix vector formulation can be used to derive practical calibration algorithms, we summarize here the main steps involved in a widely used POLSAR calibration algorithm, first derived in Quegan (1994), and then further modified by Papathanassiou (1998a) and Kimura (2004). It employs the small cross-talk hypothesis of equation (9.18) to express the relationship between calibrated and observed (uncalibrated) scattering four-element vectors in the form shown in equation (9.20).

$$\begin{bmatrix} hh\\ hv\\ vh\\ vv \end{bmatrix}_{calibrated} = Y \begin{bmatrix} k^2 & 0 & 0 & 0\\ 0 & k & 0 & 0\\ 0 & 0 & k & 0\\ 0 & 0 & 0 & 1 \end{bmatrix}^{-1} [Z]^{-1} \cdot \begin{bmatrix} hh\\ hv\\ vh\\ vv \end{bmatrix}_{uncalibrated}$$
(9.20)

Here two scalar factors $Y = r_{22}t_{22}$ $k = \frac{r_{11}}{r_{22}}$ have been factored from the matrix inverse. These are found by imaging a *single* point target with a known scattering

matrix. Often a trihedral corner reflector (Figure 1.21) is used for this purpose. It has the identity matrix as a true scattering matrix, and hence by measuring the ratio of apparent copolarised elements for the pixel we can establish a direct estimate of 'k'. The radiometric factor Y can then be determined from the known radar cross-section (RCS) of the trihedral. The 4×4 matrix $[Z]^{-1}$ can be written in terms of a set of four cross-talk ratios *u*, *v*, *w* and *z*, and a factor 'a' as shown in equation (9.21):

$$[Z]^{-1} = \frac{1}{a} \begin{bmatrix} 1 & -v & -w & vw \\ -az & a & azw & -aw \\ -u & uv & 1 & -v \\ azu & -au & -az & a \end{bmatrix}$$

$$u = \frac{r_{21}}{r_{11}} \quad v = \frac{t_{21}}{t_{22}} \quad w = \frac{r_{12}}{r_{22}} \quad z = \frac{t_{12}}{t_{11}} \quad a = \frac{r_{22}t_{11}}{r_{11}t_{22}}$$
(9.21)

Importantly, these five parameters can be estimated from observations of the covariance matrix for a distributed or depolarising region of the image. Solving for these components requires two key assumptions:

- 1. The first is reciprocity, so that for the true matrix HV = VH. Hence any departure from this in the measured data is attributed to the effects of the system distortions *u*, *v*, *w*, *z*, and *a*.
- 2. In addition, we also need to assume that the depolariser has reflection symmetry, so that the true covariance matrix has zero elements whenever co- and crosspolarised channels are multiplied (see Section 2.4.2).

The second is a more restrictive assumption, as it may not be true in the presence of surface slopes or in heterogeneous regions, where discrete point scatterers exist (in urban areas, for example). It also forces the use of a large number of looks to reduce bias in the estimation of zero cross-products. For this reason, calibration must be applied over very flat homogeneous regions containing strong depolarising effects (volume scattering). Flat, forested areas (such as the Amazon basin for spaceborne sensors) are typical of regions of choice. On the other hand, urban and mountainous regions (even if vegetated) must be avoided in the calibration process, as they are likely to violate reflection symmetry. In practice these are masked out of the calibration by first estimating the co-/crosscorrelations and rejecting pixels where this is high (Papathanassiou, 1998a; Kimura, 2004).

The detailed algorithm is shown in equation (9.22). Starting on the right we form the Quadpol scattering vector for each pixel, and then average over azimuth to obtain a 4×4 covariance matrix [*C*] (having also masked out pixels with low SNR as well as those which violate reflection symmetry). These elements can then be used to solve for all the unknowns:

$$\underline{k}_{clutter} = \begin{bmatrix} hh\\ hv\\ vh\\ vv \end{bmatrix} \Rightarrow [C] = \langle \underline{k}_{clutter} \underline{k}_{clutter}^{*T} \rangle = \begin{bmatrix} c_{11} & c_{12} & c_{13} & c_{14}\\ c_{21} & c_{22} & c_{23} & c_{24}\\ c_{31} & c_{32} & c_{33} & c_{34}\\ c_{41} & c_{42} & c_{43} & c_{44} \end{bmatrix}$$

$$\Rightarrow \begin{cases} u = \frac{c_{44}c_{31} - c_{41}c_{34}}{c_{11}c_{44} - c_{14}c_{41}} \\ v = \frac{c_{11}c_{34} - c_{31}c_{14}}{c_{11}c_{44} - c_{14}c_{41}} \\ w = \frac{c_{11}c_{24} - c_{21}c_{14}}{c_{11}c_{44} - c_{14}c_{41}} \\ z = \frac{c_{44}c_{21} - c_{41}c_{24}}{c_{11}c_{44} - c_{14}c_{41}} \\ a_{1} = \frac{c_{33} - uc_{13} - vc_{43}}{c_{23} - zc_{13} - wc_{43}} \\ a_{2} = \frac{(c_{23} - zc_{13} - wc_{43})^{*}}{c_{22} - z^{*}c_{21} - w^{*}c_{24}} \end{cases}$$
(9.22)

These calibration parameters are then used to correct the observed single look complex (SLC) data for each pixel by the matrix multiplication shown in equations (9.20) and (9.21). Note that the two parameters a_1 and a_2 are combined to estimate 'a' as shown in equation (9.23) (assuming equal noise in all polarisation channels):

$$a = \frac{|a_1a_2| - 1 + \sqrt{(|a_1a_2| - 1)^2 + 4|a_2|^2}}{2|a_2|}$$
(9.23)

From this we can symmetrize the matrix; that is, we can estimate the true crosspolarised component as a linear combination of the measured crosspolarised signals, as shown in equation (9.24):

$$xx = \frac{(a^*hv + vh)}{1 + aa^*}$$
(9.24)

Two recent modifications of this basic algorithm have been proposed. In the first—by Kimura (2004)—the assumptions about equal noise in all channels can be relaxed. This can in principle allow the treatment of low SNR regions, as can occur in power-limited applications such as spaceborne sensors. However, given the additional multi-looking requirements in noisy regions and the higher SNR achievable with airborne systems, it is often easier just to mask out those few areas of low SNR using, for example, the fourth eigenvalue of [C], or more simply the HV/VH coherence. If the HV/VH coherence is less than, say, 0.9 (around 10 dB SNR), then mask out the pixels from the calibration algorithm.

The second recent development (Ainsworth, 2006) has been the relaxing of the requirement for zero correlation between co- and crosspolarisation channels (the reflection symmetry assumption). This allows application of the technique over a much wider range of terrain types at the expense of computational complexity. (An iterative algorithm is now required where the parameters are first estimated from the data and then fed back into the model to improve the estimation.)

9.3.4 Compact polarimetry

Sometimes the complexity, bandwidth and range swath coverage restrictions of switching the transmitter polarisation are undesirable, and so-called compact

polarimetry systems have been developed as a compromise (Souyris, 2005; Raney, 2006, 2007). In these systems the transmitter polarisation state is *fixed*, but the dual channel coherent receiver configuration is maintained. This yields measurement of only part of the complex scattering matrix, although interesting permutations arise by allowing the transmitter state to be a different polarisation to that of the receiver; for example, transmitting circular polarisation and receiving horizontal and vertical linear components. In this section we outline a general approach to such compact designs and highlight some of their strengths and weaknesses.

In general we start by considering the *S* matrix represented in an arbitrary orthogonal basis *xy* used in the receiver. The fixed transmitter polarisation is then represented by complex components p_x and p_y in this basis. The key constraint of compact polarimetry is that p_x and p_y are fixed and form a unitary vector (with unit amplitude). The two orthogonal receiver channels then measure complex signals s_1 and s_2 , as shown in equation (9.25):

$$\begin{bmatrix} s_1 \\ s_2 \end{bmatrix} = \begin{bmatrix} S_{XX} & S_{XY} \\ S_{YX} & S_{YY} \end{bmatrix} \begin{bmatrix} p_x \\ p_y \end{bmatrix}$$
(9.25)

Each of these received signals is a linear combination of the elements of [S]. The real utility of compact polarimeters lies, however, not in coherent analyses but in the characterization of depolarisers. In this case, interest centres not so much on the complex signals s_1 and s_2 but on their 2×2 coherency matrix [J], as shown in equation (9.26):

$$s_{1} = S_{XX}p_{x} + S_{XY}p_{y} s_{2} = S_{YX}p_{x} + S_{YY}p_{y}$$
$$\xrightarrow{|p_{x}|^{2} + |p_{y}|^{2} = 1} [J] = \begin{bmatrix} \langle s_{1}s_{1}^{*} \rangle & \langle s_{1}s_{2}^{*} \rangle \\ \langle s_{2}s_{1}^{*} \rangle & \langle s_{2}s_{2}^{*} \rangle \end{bmatrix} = \begin{bmatrix} J_{XX} & J_{XY} \\ J_{YX} & J_{YY} \end{bmatrix}$$
(9.26)

This matrix has only four parameters, while the full scattering coherency matrix has up to sixteen. However, by assuming two symmetries in the scattering process we can reduce this discrepancy. The first—reciprocity in backscatter forces $S_{xy} = S_{yx}$ and the full coherency matrix—then has rank 3 with nine parameters. The second—reflection symmetry with an axis aligned parallel to *x* or *y*—forces cross-products involving mixed co- and crosspolar terms to zero; that is, $\langle S_{XX} S_{XY} \rangle = \langle S_{YY} S_{YX} \rangle = 0$. This reduces the scattering coherency matrix [*T*] and covariance matrix [*C*] in the *xy* basis to the reduced 3 × 3 forms shown in equation (9.27), both of which have only five unknowns.

$$[T] = \begin{bmatrix} t_{11} & t_{12} & 0\\ t_{12}^* & t_{22} & 0\\ 0 & 0 & t_{33} \end{bmatrix} \Leftrightarrow [C] = \begin{bmatrix} c_{11} & 0 & c_{13}\\ 0 & c_{22} & 0\\ c_{31}^* & 0 & c_{33} \end{bmatrix}$$
(9.27)

Now we wish to relate the four observations of [J] obtained in compact polarimetry to the five unknowns of the full scattering covariance matrix under reciprocal reflection symmetry. This we can do by expanding equation (9.26) and using reciprocity and reflection symmetry relations to obtain the following set of linear equations:

$$\begin{bmatrix} p_{x}p_{x}^{*} & p_{y}p_{y}^{*} & 0 & 0 & 0\\ 0 & p_{x}p_{x}^{*} & p_{y}p_{y}^{*} & 0 & 0\\ 0 & \operatorname{Re}(p_{x}p_{y}^{*}) & 0 & \operatorname{Re}(p_{x}p_{y}^{*}) & -\operatorname{Im}(p_{x}p_{y}^{*})\\ 0 & -\operatorname{Im}(p_{x}p_{y}^{*}) & 0 & \operatorname{Im}(p_{x}p_{y}^{*}) & \operatorname{Re}(p_{x}p_{y}^{*}) \end{bmatrix} \begin{bmatrix} c_{11} \\ c_{22} \\ c_{33} \\ \operatorname{Re}(c_{13}) \\ \operatorname{Im}(c_{13}) \end{bmatrix} = \begin{bmatrix} J_{XX} \\ J_{YY} \\ \operatorname{Re}(J_{XY}) \\ \operatorname{Im}(J_{XY}) \end{bmatrix} \Rightarrow [P]_{\underline{C}} = \underline{j}$$

$$(9.28)$$

Here we have four equations in five unknowns, and so cannot solve for all five elements of [*C*], whatever the choice of p_x and p_y . There are then three important special cases of compact polarimetry that arise in practice. They all derive from the choice of xy = HV; that is, for linear horizontal and vertical on receive. In the simplest case we can then choose $p_x = 1$, $p_y = 0$ —fixed horizontal transmit. In this case the [*P*] matrix takes the following form:

Here we see that [J] then only contains information about the scattered power in co- and cross-channels. (Remember that we are assuming reflection symmetry and so the HH and HV channels are uncorrelated.) In order to access information related to the other elements a different choice of p_x and p_y are required. In the $\pi/4$ compact mode, for example, the transmitter is set to 45° linear— $p_x = p_y = 1/\sqrt{2}$ —and the matrix [P] takes the form shown in equation (9.30):

$$p_x = \frac{1}{\sqrt{2}}, p_y = \frac{1}{\sqrt{2}} \Rightarrow [P] = \frac{1}{2} \begin{bmatrix} 1 & 1 & 0 & 0 & 0\\ 0 & 1 & 1 & 0 & 0\\ 0 & 1 & 0 & 1 & 0\\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$
(9.30)

We note two important aspects of this mode. The first is that we now have access to linear combinations of all the elements of [C], and hence some sensitivity to all the elements of the covariance matrix. The second is the factor of 1/2 in front of the matrix [P]. This implies a 3-dB loss of signal compared to a full [S] matrix system. Such signal loss is an inevitable consequence of mismatching the transmitter and receiver bases. Finally, another mode that has been proposed is to transmit circular polarisation: $p_x = 1/\sqrt{2}$, $p_y = \pm i/\sqrt{2}$. This case is very similar to the $\pi/4$ mode, but with a [P] matrix of the form shown in equation (9.31):

$$p_x = \frac{1}{\sqrt{2}}, \quad p_y = \frac{\pm i}{\sqrt{2}} \Rightarrow [P] = \frac{1}{2} \begin{bmatrix} 1 & 1 & 0 & 0 & 0\\ 0 & 1 & 1 & 0 & 0\\ 0 & 0 & 0 & 0 & \pm 1\\ 0 & \pm 1 & 0 & \mp 1 & 0 \end{bmatrix}$$
(9.31)

Again we see that there is a 3-dB loss of signal due to antenna mismatch, but again there is information from all components of [C] present in the mixture.

Some authors have tried to extend this approach so as to be able to reconstruct the reflection symmetric [C] matrix in full (Souyris, 2005). To do this we require an extra constraint equation between the elements of [C] to reduce the number of unknown to four, so matching the number of observations. Ideally we would like to find an extra linear relationship so that we could make [P] a 5×5 square matrix and then solve for the elements of <u>c</u> by matrix inversion. However, so far no such linear relationship has been found, and instead a non-linear constraint is widely used. We can motivate the development of this approach as follows.

One way to reduce the number of unknowns in [C] is to assume a model of scattering. In common with our discussions in Sections 4.2.3 and 7.4.1, we adopt a random-volume-over-ground (RVOG) model for scattering by natural terrain. In this case we assume the volume scattering component shows the much stronger azimuthal symmetry, and it is only the presence of the direct surface or dihedral returns that break this symmetry and leads to a reflection symmetric composite. This model can now be used to relate the normalized level of crosspolarisation to the copolar coherence, as shown in equation (9.32). We start by considering the limiting case of zero surface component; that is, pure volume scattering. In this case the coherency matrix is diagonal with two degenerate eigenvalues. This leads to an additional relationship between the crosspolarised power and the power in the second Pauli channel as shown. By expanding and using the fact that for azimuthal symmetry the copolarised powers in XX and YY are equal, we obtain a relationship between the HH/VV coherence and normalized crosspolarised power. The key assumption we can now make is that this relationship applies even when we add a non-zero surface component.

$$[T] = [T_{s}] + [T_{V}]$$

$$[T_{V}] = \begin{bmatrix} t_{11} & 0 & 0 \\ 0 & t_{22} & 0 \\ 0 & 0 & t_{22} \end{bmatrix} \Rightarrow 4 \langle |S_{XY}|^{2} \rangle = \langle |S_{XX} - S_{YY}|^{2} \rangle$$

$$= \langle |S_{XX}|^{2} \rangle + \langle |S_{YY}|^{2} \rangle - 2 \operatorname{Re}(S_{XX}S_{YY}^{*}) = 2 \langle |S_{XX}|^{2} \rangle - 2 \operatorname{Re}(S_{XX}S_{YY}^{*})$$

$$\Rightarrow \frac{4 \langle |S_{XY}|^{2} \rangle}{\langle |S_{XX}|^{2} \rangle + \langle |S_{YY}|^{2} \rangle} = 1 - \frac{2 \operatorname{Re}(S_{XX}S_{YY}^{*})}{\langle |S_{XX}|^{2} \rangle + \langle |S_{YY}|^{2} \rangle}$$

$$= 1 - \frac{2 \operatorname{Re}(S_{XX}S_{YY}^{*})}{2 \langle |S_{XX}|^{2} \rangle} = 1 - |\gamma_{XXYY}|$$

$$\Rightarrow \frac{\langle |S_{XY}|^{2} \rangle}{\langle |S_{XX}|^{2} \rangle + \langle |S_{YY}|^{2} \rangle} = \frac{1}{4} (1 - |\gamma_{XXYY}|) \qquad (9.32)$$

To check this we first ask what happens in the limit as the volume tends to zero and we are left with bare surface scattering. In this case (according to the RVOG model) the coherency matrix is rank-1 (or with very small secondary eigenvalues), and is therefore represented by a symmetric scattering matrix, which we also assume is diagonal in the XY basis (due to Bragg scattering

from a flat surface, for example). Hence it has zero crosspolarisation combined with a high polarimetric coherence equal to unity. We see that this combination is still consistent with equation (9.32). For the general mixed case between these two extremes we can adopt a simple two-component decomposition as shown in equation (9.33):

$$[T] = [T_s] + [T_V] = m_s \begin{bmatrix} \cos^2 \alpha & \sin \alpha \cos \alpha e^{i\delta} & 0\\ \sin \alpha \cos \alpha e^{-i\delta} & \sin^2 \alpha & 0\\ 0 & 0 & 0 \end{bmatrix} + m_v \begin{bmatrix} 0.5 & 0 & 0\\ 0 & 0.25 & 0\\ 0 & 0 & 0.25 \end{bmatrix}$$
$$\Rightarrow [C] = \frac{m_s}{2} \begin{bmatrix} 1 + \sin 2\alpha \cos \delta & 0 & \cos 2\alpha + i \sin 2\alpha \sin \delta\\ 0 & 0 & 0\\ \cos 2\alpha - i \sin 2\alpha \sin \delta & 0 & 1 - \sin 2\alpha \cos \delta \end{bmatrix} + \frac{m_v}{8} \begin{bmatrix} 3 & 0 & 1\\ 0 & 2 & 0\\ 1 & 0 & 3 \end{bmatrix}$$
(9.33)

Here we combine two components: one a rank-1 surface mechanism with magnitude m_s , and the second a random dipole cloud with scattering cross section m_v . This is very similar to the Freeman decomposition (see Section 4.2.4) or the RVOG model (see Section 7.4.1). We can now express the cross-to-copolarised scattering ratio and HH/VV coherence as functions of the surface-to-volume scattering ratio $\mu = m_s/m_v$ and scattering mechanisms α and δ , as shown in equation (9.34):

$$\frac{4\langle |S_{HV}|^2 \rangle}{\langle |S_{HH}|^2 \rangle + \langle |S_{VV}|^2 \rangle} = \frac{1}{2\mu + \frac{3}{2}}$$

$$1 - |\gamma_{HHVV}| = 1 - \left| \frac{\frac{1}{4} + \mu(\cos 2\alpha + i\sin 2\alpha \sin \delta)}{\sqrt{\mu^2(1 - \sin^2 2\alpha \cos^2 \delta) + \frac{3}{2}\mu + \frac{9}{16}}} \right|$$
(9.34)

We can now use this to check the equality of equation (9.32) for *arbitrary* mixtures of surface and volume scattering mechanisms. Figure 9.12 shows some example calculations. Here we plot, along the x axis, the cross-to-copolar ratio, and along the y axis one minus the coherence amplitude. For equation (9.32) to be valid, therefore, we require the points to lie along a line at 45° . We show how the two parameters vary for μ ranging from -30 dB to +30 dB; that is, from the limiting cases of zero surface to zero volume scattering. We show the results for steps of 15° in alpha (always with $\delta = 0$ to simplify the situation), starting from zero. We note that for $\alpha = 0$ the equality holds for all mixtures, and that the two limiting points of zero surface (the origin) and zero volume (when both approach 2/3) are also satisfied for all scattering mechanisms, as expected. However, for alpha angles greater than 30° we note significant departures from the model. In particular, for $alpha = 45^{\circ}$ we see that we have a situation where the coherence can be zero even when there is low crosspolarisation. This arises because one of HH or VV scattering coefficients goes to zero for this mechanism.

More significantly we see that for all alpha values greater than 45° for dihedral scattering of all types—there is always a μ value that leads to zero coherence, and consequently to large deviations from the simple linear



Fig. 9.12 The variation of cross-tocopolarised ratio versus 1-HHVV coherence for varying mixture of surface and volume scattering and different scattering mechanisms

relationship. We can see from equation (9.34) that this arises when $\cos 2\alpha$ is negative, as it can then cancel the positive numerator contribution from the volume scattering. As such this effect has its origin in the 180-degrees phase shift caused by double reflection.

Despite these limitations, relations such as equation (9.32) are widely used in the compact polarimetry community. The reason is that by combining this result with [P] leads to a set of five (non-linear) equations in five unknowns (the elements of [C]), as shown in equation (9.35), which can then be solved by iteration for *all* the elements of [C] and therefore [T] or the Mueller matrix [M].

$$\begin{bmatrix} p_{x}p_{x}^{*} & p_{y}p_{y}^{*} & 0 & 0 & 0\\ 0 & p_{x}p_{x}^{*} & p_{y}p_{y}^{*} & 0 & 0\\ 0 & \text{Re}(p_{x}p_{y}^{*}) & 0 & \text{Re}(p_{x}p_{y}^{*}) & -\text{Im}(p_{x}p_{y}^{*})\\ 0 & -\text{Im}(p_{x}p_{y}^{*}) & 0 & \text{Im}(p_{x}p_{y}^{*}) & \text{Re}(p_{x}p_{y}^{*}) \end{bmatrix} \begin{bmatrix} c_{11} \\ c_{22} \\ c_{33} \\ \text{Re}(c_{13}) \\ \text{Im}(c_{13}) \end{bmatrix} = \begin{bmatrix} J_{XX} \\ J_{YY} \\ \text{Re}(J_{XY}) \\ \text{Im}(J_{XY}) \end{bmatrix}$$
$$\frac{4c_{22}}{c_{11}+c_{33}} + \left| \frac{c_{13}}{\sqrt{c_{11}c_{33}}} \right| - 1 = 0$$
(9.35)

In this way, compact polarimetry can be used to provide *estimates* for the full coherency or covariance matrix elements without compromising the PRF requirements of a single channel SAR system. Note, however, that this is only true for the class of depolarisers satisfying the combined assumptions of reciprocity, reflection symmetry and especially the equality between copolar coherence and cross-to-copolarised power (which, as we have shown in Figure (9.12), is the weakest assumption). In addition to the case shown above, the compact assumptions also do not apply to general point scatterers (such as rotated dihedrals or dipoles, for example) or to scattering from sloped terrain—both of which introduce correlation between co- and cross-polarisations and

lead to a high coherence, even in the presence of significant crosspolarisation. For these more general scenarios, full [S] matrix polarimetry is required. Hence the user must be aware of exactly what kind of applications are in mind when deciding on the best mode for use in imaging radar polarimetry.

Finally, we note that in [S] or compact POLSAR imaging it is important to measure the phase as well as the amplitude of the scattered signal, and hence coherent in-phase and quadrature (IQ) detection is required. If such detectors are not available, an alternative indirect measurement of [S] can be made, based entirely on incoherent (intensity) detectors only. However, in this case a combination of four transmitter states (typically linear H, V, 45-degree, and circular polarisations) and four receiving filters are used to measure the sixteen elements of the Mueller matrix [M] (see Section 2.2). From these elements, under some restrictions (see Section 2.2), we can then estimate seven of the eight components of [S]. (Absolute phase cannot be determined by this technique, and hence interferometry cannot benefit from this approach.) Alternatively we can use this to estimate the covariance or coherency matrix directly from [M]. An extreme form of this approach is to dispense with the transmitter completely and rely on natural radiation of a scene. In this case the four-element incoherent receiver measures the Stokes vector of the scattered wave (see Section 1.3.4). Normally, the incident wave in such configurations is considered a randomly polarised wave; that is, one with zero coherence, in which case only the first column of the Mueller matrix is measured, providing access to limited information about the scattering matrix (see Section 2.6). Such an approach is widely used at optical wavelengths where direct phase measurements are not possible. There have been some examples of this approach in radar applications (Boerner, 1992; Sarabandi, 1992a), but primary interest has been in coherent imaging system applications. This brings us to consider the most general case, in which imaging polarimetry and interferometry are combined into the most flexible sensor configuration: POLInSAR.

9.4 Polarimetric SAR interferometry (POLInSAR)

The final stage in our radar imaging hierarchy is to consider imaging polarimetric interferometry or POLInSAR (Cloude, 1998; Krieger, 2005). This involves measurement of the full scattering matrix with wave space diversity for two spatial tracks separated by a baseline vector <u>b</u>. This then enables SAR imaging using Stolt interpolation and inverse Fourier transforms of each of the eight complex channels, followed by interferometry between co-registered complex images using *arbitrary* complex linear combinations based on weight vectors w_1 and w_2 , as shown schematically in Figure 9.13.

This then provides maximum flexibility in terms of combined image-based polarimetric and interferometric processing. The resolution and coverage issues are the same as for polarimetry, and the same balance as regards compact or quad polarimetry must be considered. The calibration and compact polarimetry requirements do, however, deserve special attention, as they have some important differences from standard polarimetry. We now turn to consider each of these in turn.



Fig. 9.13 Schematic of steps involved in polarimetric SAR Interferometry, or POLInSAR

9.4.1 Calibration of POLINSAR systems

In this section we consider the effect of polarimetric calibration errors on coherence estimation in polarimetric interferometry. We saw in equation (9.16) that calibration errors can be represented by a distortion matrix [Z], which multiplies the true scattering vector \underline{k} to yield the observed or measured vector. Any uncorrected distortions will then impact on the estimate of the interferometric coherency matrix and hence on estimation of coherence itself. In polarimetric interferometry we must further allow for the possibility that we have different distortion matrices Z_1 and Z_2 at the two different spatial/temporal positions across the baseline. Hence we can generate the following general distorted forms of the composite 6×6 matrix Λ_2 .

$$[\Lambda_{2}] = \begin{bmatrix} Z_{1}T_{11}Z_{1}^{*T} & Z_{1}\Omega_{12}Z_{2}^{*T} \\ Z_{2}\Omega_{12}^{*T}Z_{1}^{*T} & Z_{2}T_{22}Z_{2}^{*T} \end{bmatrix}$$

$$\Rightarrow \tilde{\gamma}(\underline{w}_{1}, \underline{w}_{2}) = \frac{\underline{w}_{1}^{*T}Z_{1}\Omega_{12}Z_{2}^{*T}\underline{w}_{2}}{\sqrt{\underline{w}_{1}^{*T}Z_{1}T_{11}Z_{1}^{*T}\underline{w}_{1}} \cdot \sqrt{\underline{w}_{2}^{*T}Z_{2}T_{22}Z_{2}^{*T}\underline{w}_{2}}$$
(9.36)

This impacts on the estimation of coherence for scattering vectors \underline{w}_1 and \underline{w}_2 , as shown on the right-hand side of equation (9.36). In simple terms it is clear, therefore, that calibration errors do change coherence, and hence will act to distort, for example, the coherence region. In particular, if $[Z_1]$ and $[Z_2]$ are different we obtain an unknown mixture of polarimetric and interferometric coherence, which will act to distort our interpretation of scattering behaviour in the pixel by mixing interferometric and polarimetric coherences. This again points to the need for good polarimetric calibration procedures, driving the matrices Z_1 and Z_2 as close as possible to the identity matrix. However, it is one advantage of the optimization approach to POLInSAR that it provides some robustness to residual calibration errors, as we now demonstrate.

Using the distorted form of Λ_2 including calibration effects, we can now rewrite the optimization eigenvalue relations (see Section 6.2) for constrained

and unconstrained optimization, as shown in equation (9.37):

$$\stackrel{\underline{w}_{1}=\underline{w}_{2}}{\longrightarrow} \left(Z_{1}T_{11}Z_{1}^{*T} + Z_{2}T_{22}Z_{2}^{*T} \right)^{-1} \left(Z_{1}\Omega_{12}Z_{2}^{*T} + Z_{2}\Omega_{12}^{*T}Z_{1}^{*T} \right) \underline{w} = \lambda_{r}\underline{w}$$

$$if \ Z_{1} = Z_{2}$$

$$(Z^{*T})^{-1}T^{-1}\Omega Z^{*T}\underline{w} = \lambda_{r}\underline{w}$$

$$\Rightarrow T^{-1}\Omega\underline{w}' = \lambda_{r}\underline{w}'$$

$$\stackrel{\underline{w}_{1}\neq\underline{w}_{2}}{\longrightarrow} \left\{ \left(Z_{2}^{*T} \right)^{-1}T_{22}^{-1}\Omega_{12}^{*T}T_{11}^{-1}\Omega_{12}Z_{2}^{*T}\underline{w}_{2} = \lambda_{opt}\underline{w}_{2} \\
\left(Z_{1}^{*T} \right)^{-1}T_{11}^{-1}\Omega_{12}T_{22}^{-1}\Omega_{12}^{*T}Z_{1}^{*T}\underline{w}_{1} = \lambda_{opt}\underline{w}_{1} \\
\Rightarrow \left\{ \begin{array}{c} T_{22}^{-1}\Omega_{12}^{*T}T_{11}^{-1}\Omega_{12}\underline{w}_{2}' = \lambda_{opt}\underline{w}_{2} \\
T_{11}^{-1}\Omega_{12}T_{22}^{-1}\Omega_{12}^{*T}\underline{w}_{1}' = \lambda_{opt}\underline{w}_{1}' \\
\end{array} \right.$$

$$(9.37)$$

In the upper part we show the effect on constrained optimization $(w_1 = w_2)$, and see that as long as $Z_1 = Z_2$ then it follows that the effects of calibration distortion can be absorbed into the eigenvectors, and that the optimum coherence values themselves (the eigenvalues) remain unchanged. In the lower part of equation (9.37) we show a much stronger form of this result for the unconstrained optimization case. Here we see that even if Z_1 and Z_2 are different, the effects of calibration can still be absorbed into the eigenvectors, and hence the optimum coherences remain unchanged. This can be important in applications where only the coherences themselves are used rather than the eigenvectors. In such cases the calibration requirements can be relaxed compared to those required for polarimetry alone.

9.4.2 **Compact POLInSAR**

By fixing the transmit polarisation, but receiving orthogonal components x and y for two sampling positions separated by a baseline B, we obtain a 4×4 polarimetric interferometric matrix, as shown in equation (9.38):

. .

. .

. ___

$$\begin{cases} s_{x}^{1} = S_{xx}^{1} p_{x} + S_{xy}^{1} p_{y} \\ s_{y}^{1} = S_{yx}^{1} p_{x} + S_{yy}^{1} p_{y} \\ s_{x}^{2} = S_{xx}^{2} p_{x} + S_{xy}^{2} p_{y} \\ s_{y}^{2} = S_{yx}^{2} p_{x} + S_{yy}^{2} p_{y} \end{cases} \xrightarrow{|p_{x}|^{2} + |p_{y}|^{2} = 1} [J] = \begin{bmatrix} \left\langle s_{x}^{1} s_{x}^{1*} \right\rangle \ \left\langle s_{x}^{1} s_{y}^{1*} \right\rangle \ \left\langle s_{y}^{1} s_{x}^{2*} \right\rangle \ \left\langle s_{y}^{2} s_{y}^{2*} \right\rangle \ \left\langle s_{y}^{2} s_{x}^{2*} \right\rangle \ \left\langle s_{y}^{2} s_{x}^{2*} \right\rangle \ \left\langle s_{y}^{2} s_{x}^{2*} \right\rangle \ \left\langle s_{y}^{2} s_{y}^{2*} \right\rangle \ \left\langle$$

This matrix can be partitioned into 2×2 sub-matrices for polarimetry and interferometry, as shown by the ^C superscript for 'Compact', and then the same optimization procedures as used for quadpol interferometry can be applied. Both constrained and unconstrained optimization algorithms can be developed.

The unconstrained optimization, for example, can be implemented as shown in equation (9.39):

$$\stackrel{\underline{w}_{1}\neq\underline{w}_{2}}{\longrightarrow} \begin{cases} T_{22}^{C-1}\Omega_{12}^{C*T}T_{11}^{C-1}\Omega_{12}^{C}\underline{w}_{2} = \lambda_{1}\lambda_{2}^{*}\underline{w}_{2} \\ T_{11}^{C-1}\Omega_{12}^{C}T_{22}^{C-1}\Omega_{12}^{C*T}\underline{w}_{1} = \lambda_{1}\lambda_{2}^{*}\underline{w}_{1} \end{cases} \lambda_{1} = \lambda_{2} = \tilde{\gamma}_{opt} \Rightarrow \begin{cases} K_{1}^{C}\underline{w}_{1} = \nu\underline{w}_{1} = \left|\gamma_{opt}\right|^{2}\underline{w}_{1} \\ K_{2}^{C}\underline{w}_{2} = \nu\underline{w}_{2} = \left|\gamma_{opt}\right|^{2}\underline{w}_{2} \end{cases} \end{cases}$$

$$(9.39)$$

We can also derive the coherence region for compact polarimetry using constrained optimization, as we did in Section 6.2. The key difference here is that we are now working in a two-dimensional rather than three-dimensional complex space, and hence the coherence region will always be an ellipse (see Section 6.2.4). It is interesting, from an applications point of view, to determine if we can use these ideas to approximate the true region shape by using such a compact system. We return to this point in Section 9.5.3, when we consider the coherence region for vegetation scattering obtained from chamber measurements.

We now turn to consider some illustrative applications of polarimetry and polarimetric interferometry.

9.5 Applications of polarimetry and interferometry

In this section we illustrate application of the theory developed in previous chapters. To do this we employ data from four sources—the first being polarimetric interferometry measurements made inside a large 10-m anechoic chamber: the European Microwave Scattering Laboratory (EMSL), located at the European Joint Research Centre (JRC), at Ispra, in Italy: http://www-emsl.jrc.it/ (Cloude, 1999; Sagues, 2000, 2001; Lopez, 2000).

The second is data from advanced computational electromagnetic simulations made by the NASA Goddard Space Center in New York: http://www. giss.nasa.gov/~crmim/ (Mishchenko, 2007). These represent an example of the very latest developments in the computer-based solution of Maxwell's equations for a complicated system of interacting particles—in this case a cluster of dielectric spheres. This approach provides a full solution free of many of the simplifying assumptions usually employed in multiple scattering calculations. One advantage of this approach is that it provides a full vector solution, including depolarisation effects, allowing us to explore full parameterization of a complex scattering problem.

The third is data from an airborne imaging POLInSAR system: the E-SAR, operated by the German Aerospace Centre (DLR) at Oberpfaffenhofen, near Munich, Germany: http://www/dlr.de/hr/en/desktopdefault.aspx. This was one of the first systems to successfully demonstrate repeat-pass polarimetric interferometry at low radar frequencies (L and P bands) (Papathanassiou, 1998; Reigber, 2000, 2001), and since then has been a major source of such data to the wider radar sciences community (Hajnsek, 2009).

Finally we employ data from the ALOS-PALSAR sensor—an L-band POL-SAR satellite system operated by the Japanese space agency JAXA, and



Fig. 9.14 Geometry of EMSL anechoic chamber in Ispra, Italy

launched in January 2006 (Rosenqvist, 2007). This spaceborne imaging radar provides a global Earth observation and monitoring role using full [S] matrix and dual polarimetry imaging modes.

9.5.1 Application 1: depolarisation by surface scattering

To illustrate the nature of depolarisation caused by rough surface scattering we employ data from the European Microwave Scattering Laboratory (EMSL): http://www-emsl.jrc.it/EMSLdata/nvt04-07-11/. The EMSL is located in a large anechoic chamber, enabling environmentally controlled broadband fully polarimetric measurements of surface and volume scattering. Figure 9.14 shows the geometry of the measurement chamber used. The transmitter is fixed, and can be used for scattering measurements at various angles of incidence θ . A separate receiver can be used for making monostatic or bistatic measurements on the same surface.

Computer-generated surface profiles with isotropic Gaussian statistics were machined for use in the surface scattering experiments. The two surfaces used are both a composite of sand + ethanediol + water, with rms heights of s = 2.5 cm (rough) and 0.4 cm (smooth), and with the same correlation length l = 6 cm. The surfaces are contained in a cylinder of 2 m diameter and 0.4 m depth, as shown in Figure 9.15. The bottom of the cylinder was lined with absorbing material to minimize boundary effects on the measurements.

The complex dielectric constant of the surface was measured experimentally, and shows some decrease with increasing frequency. To provide an idea of the values obtained, at 5 GHz the dielectric constant has a value $\varepsilon = 7 - i3$, rising to $\varepsilon = 9 - i4$ at 2 GHz, and falling to $\varepsilon = 5.5 - i2$ at 10 GHz.

Surface backscatter

Wideband scattering matrix measurements were made in monostatic mode (backscatter) over the frequency range 1–19 GHz and incidence angles θ of 10–50° (in 5 or 10-degree steps). For each angle of incidence the turntable is rotated through 360° in 5-degrees steps. In this non-imaging case, averaging



Fig. 9.15 Image of the computer manufactured rough surface located in the EMSL chamber

Fig. 9.16 Rough surface backscatter in linear basis as a function of frequency

is made over 360° of azimuth coverage (72 samples) combined with some frequency smoothing over a 160-MHz bandwidth (sixteen frequency steps). By averaging over such combined azimuth/frequency variations we ensure that the surface has a scattering coherency matrix of the reflection symmetric form shown in equation (2.75). Starting with the rough surface, Figure 9.16 shows the backscatter cross-section as a function of frequency for 30-degree angle of incidence. Here we show the cross-section in the linear basis, HH, HV VH and VV, being the diagonal elements of the covariance matrix [*C*]. We see that there is little change with frequency over the band, and that the copolarised channels HH and VV are some 10 dB greater than the crosspolarisation channels. As expected, we see that HV = VH (due to reciprocity and good polarimetric calibration of the data). Although the copolarised channels are also equal in amplitude (HH = VV), this is due to another reason: the particular scattering



Fig. 9.17 Coherency eigenvalue spectra for rough surface scattering

symmetries of this surface. By generating the coherency matrix [T] at each frequency and calculating its eigenvalues, we obtain the variations shown in Figure 9.17. Here we see a maximum eigenvalue around 3 dB larger than the linear HH or VV channels (due to the eigenvector, which in this case is close to the coherent sum $S_{HH} + S_{VV}$). The minimum eigenvalue is around –40 dB, and this represents an eigenvector of the form $S_{HV}-S_{VH}$. By reciprocity this should be exactly zero, but noise and residual calibration errors in the data give us around 40 dB of dynamic range in this dataset.

One interesting feature of Figure 9.17 is the presence of two small eigenvalues around 10 dB below the maximum. These represent the depolarisation subspace. The fact they are roughly equal illustrates the noise-like behaviour of this subspace—itself due to the symmetry of the rough surface. Secondly, the fact that it is a two-dimensional subspace means that not only crosspolarisation HV gives a small, depolarised return, but some coherent combination of copolarised channels is also depolarised (in this case $S_{HH}-S_{VV}$).

Figure 9.18 shows how this eigenvalue spectrum varies with angle of incidence (now for a fixed frequency of 10 GHz). Note that because of the roughness of the surface, the dominant eigenvalue does not vary much with angle of incidence. However, the depolarisation subspace shows significantly more variation, with the depolarised eigenvalues decreasing with increasing angle of incidence. One way to demonstrate the balance of this polarised/depolarised decomposition is to normalize the eigenvalues at each frequency by their sum and display them as probabilities, as shown in Figure 9.19. Here we see that at small angles the depolarised signal can be around 20% of the total, while at larger angles it reduces to around 2%. Hence, despite the roughness of this surface (and in a normalized sense it represents a wavenumber/rms roughness product $\beta s = 5.236$) the signal actually remains strongly polarised. This means that polarimetric phase and amplitude *ratios* remain coherent and can



Fig. 9.18 Variation of coherency eigenvalue spectra with angle of incidence

367

Fig. 9.19 Variation of normalized eigenvalue spectra with angle of incidence for rough surface scattering

be estimated from the data for the purposes, for example, of surface parameter estimation.

Another way to represent this information is to use the entropy/alpha approach. Figure 9.20 shows the results when applied to this rough surface scattering data. In the upper graph we show the entropy, which reduces with increasing angle of incidence. This is just another way of representing the angle of incidence dependence of the depolarised eigenvalues in Figure 9.19. We see that at small angles of incidence the entropy is over 0.6, while at larger angles it falls to around 0.1. However, by comparing Figures 9.19 and 9.20 we realize



Fig. 9.20 Entropy (upper) and alpha angle (lower) values for rough surface scattering as a function of angle of incidence

that an entropy of 0.6 still represents a strongly polarised signal. In the lower portion of Figure 9.20 we see the corresponding alpha angle. If the eigenvector were truly of the form (1,0,0), this should be zero. We see that experimentally it lies around 3°.

In conclusion, we have seen that rough surface backscattering represents only a weak depolariser, with an isotropic noise-like depolarisation subspace and a dominant eigenvector with scattering entropies below 0.6. However, the polarised eigenvector itself seems rather trivial: just the coherent sum of the HH and VV channels. The question is, do we ever obtain more interesting variation of eigenvectors, allowing us to use variation of the polarised ratios for parameter estimation? To answer this we turn attention to the smooth surface scattering behaviour.

The smooth surface is characterized by a smaller rms roughness *s* (although the same correlation length), and hence at a given frequency the product βs will be smaller and the surface electrically smoother. With this in mind we show, in Figure 9.21, the variation of linear basis backscatter as a function of frequency (at a 30-degree angle of incidence). Here we see two features of interest. The first is a lower backscatter level compared to the rough surface, with the crosspolarised response now around 20 dB below the copolarised. The second feature of interest is the separation of copolar coefficients at low frequencies. Here we see that VV is a few dB above HH, in qualitative agreement at least with the predictions of the small perturbation or Bragg scattering model.

Figure 9.22 shows the corresponding eigenvalue spectra of the coherency matrix. Again we note the small fourth eigenvalue, as expected from reciprocity, and also the presence again of a dominant eigenvalue, showing that again the backscatter is strongly polarised. However, now we see that the depolarised subspace is anisotropic; that is, that the second and third eigenvalues are not equal, demonstrating that the depolarisation caused by the surface is



Fig. 9.21 Smooth surface backscatter in linear basis as a function of frequency

Fig. 9.22 Coherency eigenvalue spectra for smooth surface scattering

now polarisation dependent. Again we can expose the polarised nature of the scattering by normalizing the eigenvalue spectra to have unit sum. Figure 9.23 shows the results as a function of angle of incidence (for a fixed low frequency of 2 GHz). Here we see a strongly polarised response, with the depolarised power making up less than a few percent of the scattered signal. The average phase angle between HH and VV is small—typically a few degrees, as shown in Figure 9.24—again agreeing qualitatively with predictions from the Bragg surface scattering model. However, the main new feature of interest is the change in the eigenvector parameters with incidence angle. Figure 9.25 shows how the amplitude of the Pauli components of the dominant eigenvector



Fig. 9.23 Variation of normalized eigenvalue spectra with angle of incidence for smooth surface scattering

Fig. 9.24 HH/VV phase for the smooth surface at 30-degree angle of incidence

vary across the spectrum. Here we see that while the first component is still dominant, the second Pauli component is non-zero and increases with angle of incidence. This is reflected in the corresponding entropy/alpha variation shown in Figure 9.26. Here we see low scattering entropy at all angles, but combined with an alpha parameter that steadily increases with angle of incidence. This variation is directly related to the dielectric constant of the surface, as we now demonstrate.

We begin by noting that the simple first-order Bragg surface scattering model can be used to estimate an alpha parameter, which is a function of the ratio of sum





Fig. 9.26 Entropy (upper) and alpha angle (lower) values at 2 GHz for smooth surface scattering as a function of angle of incidence

and difference of copolarised scattering coefficients, and hence for a given θ depends only on the dielectric constant and not surface roughness (see Section 3.1.3). We can therefore use the Bragg model with the measured dielectric constant of the surface to predict an alpha angle variation, and compare this directly with the estimates obtained from the coherency matrix. Before doing this, however, we note that the simple Bragg model ignores any influence of wave depolarisation, which while weak, still occurs even for this smooth



Fig. 9.27 Alpha angle variation at 2 GHz for smooth surface scattering as a function of angle of incidence. Solid line is the eigenvector estimate, dotted line is the prediction of the Bragg scattering model, and dashed line is the corrected alpha value using the X-Bragg scattering model of depolarisation

surface backscatter. We saw in Section 3.2 how the extended or X-Bragg model provides a method for parameterising the effects of depolarisation on the Bragg model.

According to this approach (see equation (3.40)), the alpha parameter must be corrected for depolarisation by estimating it—not directly from one eigenvector, but from a ratio of diagonal terms of the full coherency matrix, as shown in equation (9.40):

$$\alpha_b = \tan^{-1}\left(\sqrt{\frac{t_{22} + t_{33}}{t_{11}}}\right) \tag{9.40}$$

Figure 9.27 shows how these three estimates compare for low-frequency (2-GHz scattering). In the solid line we again show the eigenvector estimates from the lower part of Figure 9.26. We show too the reference Bragg values obtained from the dielectric constant. Finally, in dash we show the corrected estimated alpha values using the X-Bragg model. We see that the correction is of the order of 2°, and for angles of incidence greater than 20° the correction brings the estimates into close agreement with the Bragg predictions. However, for angles less than 20°, both methods seem to overestimate alpha. This can be traced to the small separation of HH and VV scattering coefficients for angles near normal incidence, which makes estimation of small alpha values from experimental data more difficult.

The importance of this depolarisation correction can be made even more apparent by considering a higher frequency. At 10 GHz, for example, the surface roughness leads to $\beta s = 0.84$, which is well outside the usual bounds for validity of the simple Bragg model. Figure 9.28 shows the various alpha estimates for this high-frequency case. Here we see that the maximum eigenvector dramatically underestimates the alpha parameter, especially at high incidence angles, whereas when we add the correction for depolarisation (dashed line)



Fig. 9.28 Alpha angle variation at 10 GHz for smooth surface scattering as a function of angle of incidence. Solid line is the eigenvector estimate, dotted line is the prediction of the Bragg scattering model, and dashed line is the corrected alpha value using the X-Bragg scattering model of depolarisation

we see much better agreement. This result can, for example, be used to estimate surface dielectric constant by first using the X-Bragg model to estimate alpha corrected for surface depolarisation. From alpha, and knowing the angle of incidence, we can then obtain a direct estimate of the dielectric constant by using the standard Bragg relations.

In this section we have seen that surface backscattering provides a polarised return over a wide range of roughness scales and angles of incidence. We now turn to consider the generalization of these ideas to bistatic surface scattering—the situation in which transmitter and receiver are separated.

Bistatic surface scattering

The geometry we now consider is shown in Figure 9.29. The rough surface is illuminated by the transmitter at a fixed angle of -40° . The receiver is then moved in 10-degree steps from near backscatter to beyond specular reflection $(at +40^{\circ})$. In this case we can no longer assume that HV = VH by reciprocity, and hence must employ a full 4×4 coherency matrix formulation. It is one of our objectives here to see how important this new fourth eigenvalue and associated eigenvector is for bistatic surface scattering. Figure 9.30 shows an example: the normalized eigenvalue spectra variation with scattering angle for the rough surface at 10 GHz (β s = 5). Here we see that there is again a dominant eigenvalue and hence a strongly polarised scattering component for all angles. The second and third eigenvalues are again less than 10% of the signal, while the new fourth eigenvector, although larger than in the backscatter case, remains two orders of magnitude below the main polarised response. Note the dip in the level of depolarisation around the specular direction. This is expected, as for this scattering angle the return is more coherent due to specular reflection at the rough surface. However, we see that even away from these angles the surface return remains strongly polarised.



Fig. 9.29 Geometry of bistatic surface scattering measurements in the EMSL chamber (* = TX, and O = RX)

Fig. 9.30 Normalized eigenvalue spectra for bistatic scattering from rough surface at 10 GHz

The Pauli components of the dominant eigenvector are shown in Figure 9.31. For angles close to backscatter we again see a dominant first component, corresponding to alpha around zero. However, as the scattering angle increases so the second Pauli component becomes more important. We note that the third and fourth remain equal and close to zero over the full range of angles. Figure 9.32 shows the corresponding entropy and alpha variations. Note that the entropy is now defined in base 4, to account for the fourth eigenvalue. We see a dip in



Fig. 9.31 Pauli components of dominant eigenvector of bistatic scattering coherency matrix for rough surface at 10 GHz

Fig. 9.32 Bistatic entropy (upper) and alpha angle (lower) for bistatic

the entropy for specular scattering. The alpha angle is taken from the dominant eigenvector, and shows a steady increase with scattering angle, with a trend similar to that found in Figure 9.28. However, in this case the physical origin of this variation is quite different. Here we must use the bistatic geometry and dielectric constant in the BRDF model of Section 3.2.2 (Priest, 2000). This model can be used to predict the variation of alpha for the bistatic scattering matrix. In Figure 9.33 we show the amplitude of the Pauli components of this matrix as a function of scattering angle. Superimposed, we again show the Pauli



Fig. 9.33 Variation of Pauli coefficients for the BRDF model for the rough surface (solid lines) versus the variation of the dominant eigenvector components from Figure 9.31

components for the dominant eigenvector. We see that the simple BRDF model is quite good at explaining the trend of the alpha variation.

In conclusion, we have seen from our analysis of rough and smooth surface scattering data that bare surface scattering is characterized by a strong polarised return (the entropy is seldom above 0.5), maintained over a wide range of angles of incidence and frequencies. The corresponding dominant eigenvector is characterized by an alpha parameter in the range $0 \le \alpha \le \pi/4$, and which shows dependence on local scattering geometry (angle of incidence and bistatic scattering angle) and surface dielectric constant. We now turn to consider a different class of interactions: volume scattering, where scattering entropies and hence depolarisation can be much higher.

9.5.2 Application 2: depolarisation by volume scattering

Depolarisation by volumes can be caused by two basic physical processes: particle anisotropy (in shape or material composition), or multiple scattering between particles. We begin by looking at the latter and by considering a volume made up of wavelength-sized spherical particles. Spheres have a strong symmetry, which means that in single scattering they have zero depolarisation and hence zero scattering entropy. Regardless of size and dielectric constant, in backscatter they always yield a scattering matrix proportional to the identity (in the BSA coordinate system). However, when several such particles are brought together in a volume their mutual interactions destroy this simple picture and lead to depolarisation. To illustrate this phenomenon we consider scattering by a random cloud of dielectric spheres. We make use of some recent 'exact' calculations of scattering by a cloud of particles using the superposition T-matrix method (Mishchenko, 2000). These simulations were provided courtesy of Michael Mishchenko and his group at the NASA Goddard Institute in New York.

The three-dimensional simulations are based on numerical solution of Maxwell's equations for the cloud of particles considered as a whole (Mishchenko, 2007). In this regard there are no approximations involved, and the technique can be considered 'exact' and to include all effects of multiple and single scattering such as coherent speckle, wave depolarisation, enhanced backscattering, and so on (van Albada, 1988; Mishchenko, 1992; Macintosh, 1999). The output from the simulator is all sixteen elements of the Stokes phase matrix (which is related to the Mueller matrix (Hovenier, 2004) in the FSA or wave-based coordinate system. These can then be converted in a 1–1 mapping into the elements of the 4×4 Hermitian coherency matrix (see Section 2.3), which can then be expressed in terms of its eigenvalue decomposition. These matrix elements can be determined in the simulator for arbitrary incident and scattered wave directions. From the symmetry of the problem it is sufficient for us to consider some fixed but arbitrary incident vector and variation of the scattering angle ψ in a plane formed by the incident and scattered vectors, varying from 0° (forward scatter) to 180° (backscatter).

As an example we choose to consider scattering by lossless particles with a low dielectric constant, so that the refractive index n = 1.32 ($\varepsilon_r = 1.74$) (representing the properties of water and ice particles at visible wavelengths) and size $\beta r = 4$, where β is the wavenumber in the surrounding medium and r is the particle radius. The N particles, where N = 1...240, are randomly placed in a spherical volume of size $\beta R = 40$ (see Figure 9.34), corresponding to variations from 0.1%-24% in particle concentrations.

In this way the effect of multiple scattering can be investigated by looking at the transition from single to multiple particle configurations. By symmetry, the Mueller and coherency matrices for arbitrary scattering angle must then have the form shown in equation (9.41) (see equation 2.108):

$$[M] = \begin{bmatrix} m_{11}(\psi) & m_{12}(\psi) & 0 & 0 \\ m_{12}(\psi) & m_{22}(\psi) & 0 & 0 \\ 0 & 0 & m_{33}(\psi) & m_{34}(\psi) \\ 0 & 0 & -m_{34}(\psi) & m_{44}(\psi) \end{bmatrix} \Leftrightarrow [T] = \begin{bmatrix} t_{11} & t_{12} & 0 & 0 \\ t_{12}^* & t_{22} & 0 & 0 \\ 0 & 0 & t_{33} & 0 \\ 0 & 0 & 0 & t_{44} \end{bmatrix}$$
$$t_{11} = \frac{1}{2}(m_{11} + m_{22} + m_{33} + m_{44})$$
$$t_{12} = m_{12} - im_{34}$$
(9.41)
$$t_{22} = \frac{1}{2}(m_{11} + m_{22} - m_{33} - m_{44})$$
$$t_{33} = \frac{1}{2}(m_{11} - m_{22} + m_{33} - m_{44})$$
$$t_{44} = \frac{1}{2}(m_{11} - m_{22} - m_{33} + m_{44})$$

Fig. 9.34 Geometry for calculation of scattering by random cloud of spheres

Hence we can in general have four non-zero eigenvalues, but the eigenvectors are limited in their structure to C2—a two-dimensional complex subspace. Hence we can use the scattering sphere concept (see Section 2.4.3.2) to represent variations in the eigenvector information. We choose to consider three special cases: a single particle for reference, a low concentration of 0.5%, and a high concentration of 16%. The first parameter of interest is the phase function. This



Fig. 9.35 Phase function for three special cases: single particle, five particles, and 160 particles

is just the m_{11} element of the Stokes scattering matrix, normalized so that

$$\frac{1}{2} \int_{0}^{\pi} m_{11}(\psi) \sin \psi d\psi = 1$$
(9.42)

Equivalently, m_{11} represents one half the trace of the coherency matrix (or sum of the eigenvalues of [*T*]), which physically represents the total scattered power of the signal. Figure 9.35 shows how this function varies for three different particle concentrations (scattering angle = 0 is forward scatter in this notation). We immediately see that the particles scatter more in the forward than the backward direction (a typical feature of wavelength size particles), and that as more particles are added there is coherent addition in the forward direction, and also the appearance of a small enhanced backscatter peak in the backscatter direction (scattering angle equal to 180°). We see that the phase function becomes smoother, compared to the single particle case, as we add more particles to the volume. This is the effect of multiple scattering.

We are now particularly interested in the depolarisation properties of this process. The single particle case is trivial to consider, as it yields a single non-zero eigenvalue for all scattering angles—the eigenvector corresponding to which is obtained from the scattering matrix [S] for the particle (and which can be obtained exactly using standard Mie scattering theory; see Section 3.3.2). Of more interest are the low and high concentration results. Figures 9.36 and 9.37 show the normalized coherency spectra as a function of scattering angle for these two cases.

In the low concentration example we see a dominant eigenvalue, with the secondary eigenvalues less than 1% of the scattered power, except for a couple of peaks around 60° and 100° . In forward scatter the entropy is zero (a coherent wave propagating through the volume), and in backscatter one of the minor eigenvalues goes to zero (because of reciprocity), while the other two are equal



Fig. 9.36 Normalized coherency spectra for low concentration of particles

Fig. 9.37 Normalized coherency spectra for high concentration of particles

and lie around the 1% mark. The denser concentration starts to lose structure, as multiple scattering dominates. Here we see a smoother variation of eigenvalues, still with zero entropy in forward scatter, but much higher entropy in the backscatter hemisphere. Note again that for exact backscatter one of the eigenvalues still falls to zero as a result of reciprocity. These results can be expressed in terms of the scattering entropy, as shown in Figure 9.38. Here we show the variation of entropy for the three cases. The single particle case yields zero entropy for all scattering angles, while the denser concentration leads to more depolarisation and increased entropy, especially in the backscattering hemisphere.



Fig. 9.38 Scattering entropy of a cloud of spheres versus scattering angle for three particle concentrations

We see that the entropy peaks around 0.8 for the denser concentration but then reduces slightly for backscatter. However, if we ignore the fourth eigenvalue in the backscatter direction (because reciprocity, rather than scattering symmetries, forces it to zero) and calculate the monostatic entropy, we find that the backscatter entropy rises from 0.75 to 0.94. This is close to the maximum entropy obtained from single scattering by a cloud of prolate particle shapes, even though we are dealing with a cloud of spheres. In this way we see that multiple scattering can be an important source of depolarisation.

Turning now to the dominant eigenvector, we start by noting that it has only two non-zero elements: the first and second Pauli components. Figure 9.39 shows how these two vary in amplitude for each of the three cases considered: single particle, five particles, and 160 particles. We see that in forward scattering we have the identity matrix (first Pauli component) as consistent with forward scattering in the FSA system. In backscatter, by contrast, we have the second Pauli component with the first falling to zero. This corresponds to a backscattering matrix with equal amplitudes but 180-degree phase shift between HH and VV, as expected in the FSA system. Between these two extremes we see that the eigenvectors change significantly. Indeed, these changes not only occur in amplitude but in phase. As the eigenvector is an element of C2-the twodimensional complex space-we can map it in both amplitude and phase on the surface of the scattering sphere. This allows us to visualize changes in the complex nature of the eigenvector between forward and back scattering. Figure 9.40 shows the dominant eigenvector variation on the scattering sphere for the three cases considered. In forward scattering they all begin on the equator at zero phase. As scattering angle changes we then see combined amplitude and phase variations trace a complicated loci across the sphere, ending for backscatter on the equator again at 180°. For the single-particle case (the black line) these are just the polarimetric phase and amplitude variations due to Mie scattering by a lossless dielectric sphere. In the low concentration case we see that the



Fig. 9.39 Variation of dominant eigenvector Pauli components with scattering angle for 1, 5, and 160 particle clouds

Fig. 9.40 Alpha sphere representation of the eigenvector amplitude and phase information for N = 1 particle, N = 5 particles, and N = 160 particles

dominant eigenvector closely follows this loci, maintaining information about the particle. In the high concentration case, however, we see that the dominant eigenvector is smoothed out by the multiple scattering effects.

Another way to present these ideas is to use the alpha angle. The alpha variation corresponding to the eigenvector fluctuations in Figure 9.39 are shown in Figure 9.41. Here we see that for forward scatter $\alpha = 0$ (in the FSA coordinate system), and for backscatter this rises to $\pi/2$. The single particle and low



Fig. 9.41 Alpha angle variation of dominant eigenvector as a function of scattering angle for 1, 5, and 160 particle clouds

Fig. 9.42 Schematic of the maize scattering measurements in the EMSL

concentration cases move between these two boundaries in a series of steps, while the high concentration smoothes out these details.

In this example we have seen how multiple scattering can lead to high entropies, even for spherical particles. In the next we give a second example of volume scattering—this time from EMSL chamber measurements of a vegetated surface, where both multiple scattering and anisotropic particle shape contribute to wave depolarisation. We will then use this example of a complex depolarising problem as a test case for the application of polarimetric interferometry and polarisation coherence tomography techniques.



Fig. 9.43 Image of maize sample used for EMSL measurements

9.5.3 Application 3: coherent scattering from vegetation

In this section we consider analysis of a measurement dataset that combines surface and anisotropic volume scattering. The backscatter measurement geometry inside the EMSL anechoic chamber is shown in Figure 9.42. The target, in this case, is a sample of 6×6 maize plants of 1.8 m height, uniformly planted in a square container of sides 2 m (http://www-emsl.jrc.it/EMSLdata/). Separation between plants is around 30–35 cm. The plants are characterized by green vertical stems of around 4 cm diameter—each carrying wide leaves from a height of 40 cm to the top. The leaves themselves are 30–40 cm long and 7–8 cm wide. The leaves are oriented in a range of angles from 20–45°, as shown in Figure 9.43.

The vegetation sample is placed on a rotating turntable so that measurements can be made over 360° of azimuth for a given angle of incidence. The antenna beamwidth is such that the sample is uniformly illuminated by the transmitter at all times. In this experiment there were 72 azimuth steps of 5°. At each position the frequency is stepped across the frequency range 1.5-9.5 GHz (in 10-MHz steps), and the elevation angle is incremented in 0.25-degree steps from 44° to 45° . The complete scattering matrix [S] is measured at each frequency across the full band. In this way, multi-baseline polarimetric interferometric analysis, and even coherence tomography, can be performed, with a minimum angular baseline of $\Delta \theta = 0.25^{\circ}$ and a maximum of 5°. Note, finally, that the focus for the chamber (zero phase position for interferometry) is located around 38 cm above the soil surface of the sample. Hence the interferometric phase of the true surface position ϕ_o will not be zero, and will change with frequency and baseline. This dataset provides an interesting testbed for polarimetric, interferometric, and tomographic processing (Sagues, 2000; Lopez-Sanchez, 2006, 2007; Cloude, 2007a).

Depolarisation in vegetation scattering

We begin with an assessment of the level of depolarisation caused by this scattering environment. Figure 9.44 shows the variation of normalized eigenvalues with frequency. We note that the maximum eigenvalue is now much reduced


Fig. 9.44 Variation of normalized eigenvalue spectra with frequency for the maize sample

Fig. 9.45 Variation of Pauli components of the maximum eigenvector (maize sample)

to around 60–70% of the signal energy. The depolarisation is quite large across the spectrum, and shows some anisotropy at low frequencies, which reduces as the frequency increases.

Note that here we have increased the averaging window to 320 MHz (in addition to averaging over the 72 azimuth samples) in order to reduce the speckle noise on the estimates. The Pauli components of the maximum eigenvector vary with frequency, as shown in Figure 9.45. Here we see that at low frequencies the second Pauli component is significant, but that as frequency increases the



Fig. 9.46 Entropy (upper) and alpha (lower) variation with frequency for maize sample

Fig. 9.47 Entropy/alpha plane representation of maize data points superimposed on prolate and oblate H/α loci (see Figure 3.29)

eigenvector tends to the first Pauli component. These results all indicate a high level of wave depolarisation by the vegetation sample. This is confirmed in Figure 9.46, which shows the entropy/alpha variations corresponding to Figures 9.44 and 9.45. Here we see high entropy, around 0.8 at low frequency and falling slightly to 0.7 at higher frequencies. The dominant eigenvector alpha falls from around 20° at low frequencies to nearly zero at high frequencies. However,



Fig. 9.48 Zoom on entropy/alpha plane representation of maize data points superimposed on prolate H/α loci

given the relatively high entropy we can expect the maximum eigenvector to be corrupted by depolarisation (as we saw in the cloud of spheres in Figure 9.38, for example). Hence we again need some way of compensating the eigenvector information for the depolarisation that is occurring. One way to do this is to employ the *average* alpha, formed by the sum of the alpha values for each eigenvector weighted by the normalized eigenvalues (see Section 2.4.2.4). This leads to a representation on the entropy/alpha plane as shown in Figure 9.47, with a zoom on the data points shown in Figure 9.48. Also superimposed on these diagrams are the entropy/alpha loci for a cloud of anisotropic particles, as developed in Section 3.4.1. We show results for prolate particle clouds of varying orientation structure, and the same for oblate particle clouds. One way to interpret the maize data points is in terms of the response for an equivalent cloud of such particles. In this regard we see that at low frequencies the response is equivalent to a cloud of strongly prolate particles (with a shape anisotropy factor A_p around 10 dB), but with a non-random orientation distribution; while as frequency increases two things happen: initially (up to 6 GHz) the cloud becomes more random, while maintaining the same particle anisotropy. From 6-9 GHz, however, the behaviour changes, the effective particle shape anisotropy (A_p) reduces (which causes a drop in the entropy), and the distribution stays random (the lower curve of the entropy/alpha diagram represents azimuthal symmetry in the volume).

However, this interpretation—in terms of the response of an equivalent cloud of anisotropic particles—involves many simplifying assumptions, one of the most important being that we see only volume scattering from the maize and that the underlying surface plays no part in the scattered return. This is not necessarily a good assumption, especially at low frequencies. We can see, for example, that there is a change in scattering behaviour around 6 GHz, and cannot dismiss the idea that for frequencies lower than this the surface response is playing an important role.

To resolve these issues we could try using one of the many model-based decomposition methods based on mixed surface and volume scattering, but these require us to make even more stringent assumptions about the nature of the volume scattering. (The Freeman–Durden, for example, assumes a cloud of dipole particles.) These assumptions are not supported for this example, and so would not necessarily help us resolve issues about the true ratio of surface-to-volume scattering. Instead we choose to look at the role that radar interferometry can play in helping resolve this by adding additional information about the structure of such complex scattering problems.

9.5.4 Application 4: tomographic imaging of vegetated surfaces

Interferometric processing of wide-band signals from the EMSL chamber starts with the two calibrated complex signals $s_{1,2}$ from angular positions θ and $\theta + \Delta \theta$ at frequency *f* and with transmit/receive polarisation 'pq', as shown in equation (9.43):

$$s_1 = E_{\theta}^{pq}(f) \quad s_2 = E_{\theta + \Delta\theta}^{pq}(f) \tag{9.43}$$

The wide-band interferogram is then formed from the product of common spectral band filtered and phase offset signals, as shown in equation (9.44), where the phase offset is required because the chamber focus lies below the surface of the sample container:

$$s_1.s_2^* = E_{\theta}^{pq}(f) \cdot conj\left(E_{\theta+\Delta\theta}^{pq}\left(f - \frac{f\,\Delta\theta}{\tan\theta}\right)\right)e^{-i\frac{4\pi f\,\Delta\theta}{c\,\sin\theta}z_o}$$
(9.44)

Finally, the complex coherence for polarisation combination 'pq' and frequency f is calculated as shown in equation (9.45):

$$\tilde{\gamma}_{pq}\left(f\right) = \frac{\left\langle s_{1}s_{2}^{*}\right\rangle}{\sqrt{\left\langle s_{1}s_{1}^{*}\right\rangle \left\langle s_{2}s_{2}^{*}\right\rangle}}$$
(9.45)

Figure 9.49 shows how this coherence amplitude and phase vary for three polarisations as a function of frequency for the smallest baseline (0.25°) . We see that the coherence falls as frequency increases. This is caused by volume

decorrelation, as discussed in Section 5.2.4. To confirm this, in dash we show the coherence values expected from a simple uniform structure function (the SINC model). We also show coherence results for three polarisation channels, HH, HV and VV. We note that the SINC gives the same trend as seen in the data (in both amplitude and phase); but there are some important differences, relating, as we shall see, to departures of the vertical structure function from uniform. To further emphasize these differences we show, in Figure 9.50, coherence results for a larger baseline (0.5°) . We note the following features:

- 1. The phase centre generally lies below the SINC level (which is set at half the top height of the vegetation). This implies that the phase centre is closer to the surface of the sample than expected from volumeonly scattering models, which again implies the presence of a mixed surface/volume scattering scenario.
- 2. There is separation of the polarisation channels, with VV closer to the SINC phase level (higher in the volume) than HH. However, the phase separation between polarisations is not maximized by this special selection of states. To see how a larger separation can be achieved, we show, in Figure 9.51, the phase and amplitude variation for the 0.5-degree baseline but for two of the Pauli channels. In particular we again show the crosspolar HV channel, but now compare it with the coherent combination of a difference of copolar channels HH–VV. Here we see much better phase separation across the band. This fits a physical interpretation based on a mixed surface/volume scenario with the effective surface component, in fact caused by double bounce or dihedral scattering, which introduces a 180-degree polarimetric phase shift between copolar channels.



Fig. 9.49 Variation of interferometric coherence amplitude (upper) and phase (lower) for 0.25-degree baseline and for polarisations VV and VH



Fig. 9.50 Variation of interferometric coherence amplitude (upper) and phase (lower) for 0.5-degree baseline, and for polarisations VV and VH

To confirm this we can use these complex coherences to estimate the vertical structure function using coherence tomography (Cloude, 2007a; Zhou, 2006). In this case we know from the chamber geometry the true phase of the surface position and the height of the vegetation, and so can use these directly to reconstruct the profile using the baseline geometry. Figure 9.52 shows how the interferometric wavenumber β_z varies across the spectrum for the two baselines. This can then be used with the known vegetation height (1.8 m) to calculate the normalization parameter k_v , as shown in Figure 9.53. We see that it varies from around 0.4 for the small baseline at low frequencies to more than 4.0 for the larger baseline at high frequency. These values provide good sensitivity to changes in structure function in the vegetation layer (see Section 8.4). These can then be used with the known surface phase to reconstruct a truncated Legendre series expansion of the structure function. With one baseline we obtain a second-order expansion, while by combining data from two baselines we obtain a fourth-order reconstruction. Figures 9.54 and 9.55 show the reconstructed vertical profiles through the vegetation as a function of frequency for the single and dual baseline data. We note the following important features:

- 1. The dual baseline data has higher vertical resolution than the single. This is due to the presence of higher-order terms in the Legendre series reconstruction.
- 2. Both single and dual baseline datasets show a strong surface component at low frequencies (below 6 GHz), with a more dominant volume contribution for frequencies above this. However, the volume component is not a simple exponential function. It shows a small response from the top with a peak near the centre of the layer. This therefore represents a



Fig. 9.51 Variation of interferometric coherence amplitude (upper) and phase (lower) for 0.5-degree baseline, and for two polarisations HH-VV and HV

Fig. 9.52 Interferometric wavenumber as a function of frequency for the two baselines used in the EMSL chamber

simple example where the RVOG model assumptions do not seem to be valid (see Section 7.4).

3. We note important differences between polarisations. In particular we see that VV has a much lower surface response at higher frequencies than HH. The idea that this arises from a double bounce contribution can be confirmed by imaging in the HH–VV Pauli channel, as shown in



Fig. 9.53 Normalized wavenumber/height product factor k_v as a function of frequency for the two baselines used in the EMSL chamber

Fig. 9.54 Reconstructed single-baseline vertical profiles for the 1.8-m vegetation layer in HH (upper), HV (centre), and VV (lower) polarisations

Figure 9.56 for the single (upper) and dual (lower) baseline data. Here we see a strong surface component across most of the spectrum, confirming that the structure function in this Pauli channel remains localized around the surface.

In conclusion, we have seen that the presence of vegetation on a surface causes complexity in backscatter, with an increase in the scattering entropy and hence in the level of wave depolarisation. This causes a drop in polarimetric coherence and hence an erosion of our ability to exploit polarimetric phase information.



Fig. 9.55 Reconstructed dual-baseline vertical profiles for the 1.8-m vegetation layer in HH (upper), HV (centre), and VV (lower) polarisations

Fig. 9.56 Reconstructed vertical profiles for the 1.8-m vegetation layer in the HH-VV channel for single (upper), and dual (lower) baselines

However, by combining polarimetry with interferometry we can control the variation of interferometric coherence with polarisation and use this to quantify more carefully the balance of surface and volume scattering.

We now turn to an example that combines all these ideas, but in a more challenging environment more typical of remote sensing applications: airborne radar imaging.

9.5.5 Application 5: forest height estimation using POLInSAR

E-SAR is an airborne multi-frequency SAR system operated by the German Aerospace centre (DLR) (Papathanassiou, 1998b; Hajnsek, 2009). It operates in four frequency bands—X (9.6 GHz), C (5.5 GHz), L (1.3 GHz), and P (350 MHz)—with a repeat-pass quad polarimetric interferometry mode at the two lower bands of L and P. It operates with a range swath width of 3–5 km, with range and azimuth resolutions of the order of 1.5×1.5 m, so providing multichannel images for data analysis. In this section we consider use of the lowest-frequency P-band in POLInSAR mode for forest height estimation in tropical forests.

We employ P-band repeat-pass polarimetric data collected by the DLR E-SAR airborne system as part of the ESA sponsored INDREX-II campaign (November 2004) (Hajnsek, 2009). We concentrate on tracks collected over the Mawas forest reserve in central Kalimantan (114° 36' E, 2° 12' S). This site is a typical example of a tropical peat swamp forest environment. The forest covers a large area (540,000 ha), and has a wide biomass range: 50-400 ton/ha, with a corresponding height range of 5-30 m. One key objective of this study was to investigate the potential of POLInSAR to retrieve forest biomass from height using allometric relationships (Mette, 2004, 2007; Woodhouse, 2006). The more traditional route to biomass follows directly from radar backscatter (Imhoff, 1995), but this approach is plagued by two key issues: backscatter saturation-especially at high frequencies-and high variability due to sensitivity to changes in forest structure (Woodhouse, 2006). Height estimates, on the other hand, overcome such saturation effects and, by using a robust height estimation algorithm, can be made tolerant to changes in structure (Mette, 2004; Treuhaft, 2004; Papathanassiou, 2005; Praks, 2007).

Here we make use of two tracks at P-band ($\lambda = 0.86$ m), with a nominal horizontal baseline of 30 m (and 75 minutes time baseline). The single-look complex (SLC) SAR images are provided with high resolution, operating with a slant range/azimuth pixel size of 1.4985 m/0.72 m at P-band. One key advantage of this site is the exceptionally flat topography, with slopes less than 0.1%. A further reason for employing data for the Mawas-E site is the availability of *in situ* tree-height measurements for validation. In all, 1,049 trees were measured in two parallel transects. Figure 9.57 shows a radar backscatter image of the scene with the *in situ* transects marked as black lines. In the upper picture we show an aerial photograph of the scene. We note the access road and nonforested area to the right, with the main forest to the left. In the lower part we show a P-band polarised power (maximum eigenvalue of [*T*]) image. To visualize the polarimetric information we show an entropy image in Figure 9.58. Here we see high entropy over the forest, and lower entropy (more polarised signal) over the non-forested region to the right.

Turning now to interferometry, Figure 9.59 shows a raw interferogram (before flat Earth removal) of the scene (in HH polarisation). Note that the presence of the forest can be clearly seen as an increase in phase variance. This can be confirmed by calculating the corresponding coherence. Figure 9.60 shows the interferometric coherence as a greyscale image with white = 1 and black = 0. Note the high (white) coherence over the non-forested region, with the forest showing a range-dependent variation increasing from near to far



Fig. 9.57 Aerial photograph of Mawas-E test area (top), and P band POLSAR Image (lower), with two *in situ* data transects marked as black lines

Fig. 9.58 Entropy image of MAWAS P-band data



Fig. 9.59 P-band HH repeat pass interferometric phase for forest/non-forest boundary



Fig. 9.60 Interferometric coherence of the MAWAS-E test area for P-band HH channel

range. This is characteristic of volume decorrelation, with the 30-m horizontal baseline having a higher effective β_z value in the near range (see equation (5.15)). This leads to more decorrelation for a given tree height at near range than far. We can now combine the coherence phase and amplitude information to estimate surface topography and top height, using the coherence separation optimization of equation (6.29).

In Figure 9.61 we show an image of the highest P-band phase centre (upper) and a transect through each of the two *in situ* datasets (lower). Here we show the location of the two optimum phase centres, noting around 5 m of separation between the optima. In solid, the surface topography estimate from a line fit; and in dash, the top height estimates from equation (8.38). We also show data points corresponding to *in situ* tree-height measurements (located around 800 m in azimuth). We note good general agreement with the POLInSAR height estimates, with around 20 m tree-height in the upper, and only 10 m in the lower transect.

These results indicate how the variation of interferometric coherence with polarisation can be used to estimate important biophysical parameters of interest, such as forest height, with key implications for the estimate of biomass. One key question remains as to whether we can now scale these ideas up to continental or global scale using satellite remote sensing technology. We now turn to consider this final topic in more detail.

9.5.6 Application 6: spaceborne satellite radar polarimetry

Finally, we turn to consider data from the first fully polarimetric radar satellite system to be launched: the Phased Array L-band SAR (PALSAR) on board the Advanced Land Observation Satellite (ALOS), launched by the Japanese Space Agency (JAXA) in January 2006 (Rosenqvist, 2007). Figure 9.62 shows an image of the satellite, with the large solar panels on the right and the SAR



Fig. 9.61 P-band phase centre separations along upper and lower azimuth transects for optimum polarizations, showing *in situ* height measurements as stars

Parameter	Value	
Wavelength	0.236 m	
Launch date	Jan 24 th 2006	
Orbit height	691 km	
Orbit Repeat	46 days	
Chirp bandwidth	14 MHz	
Peak transmit power	2 kW	
Duty cycle	3,5% (7 %/2)	
Noise figure	4 dB	
PRF (Quadpol)	3.8 kHz	
Antenna size (Tx, Rx)	8.9 m × 3.1 m	
Quadpol mode incidence angle	21.5 degrees	



Fig. 9.62 Schematic image and technical details of the ALOS-PALSAR satellite radar system

antenna located on the left. Some key system parameters are also shown on the right-hand side.

We note four key points in particular. The first is that the transmitted bandwidth in polarimetric mode is only 14 MHz, which is much smaller than the airborne E-SAR system, and hence the range resolution is poorer (around 11 m





Fig. 9.63 Satellite SAR Images (HH left, and HV right) for a trihedral corner reflector before Faraday rotation correction

slant range for PALSAR compared to 1.5 m for E-SAR). The second point is the relatively large PRF in polarimetric mode (around 4 kHz), which as mentioned in Section 9.3, limits the range coverage of the sensor. In this case the range swath width is reduced to around 15 km in slant range, which for 21.5degree incidence (the default mode for the satellite) translates to around 40 km of ground range. Thirdly, we note that the satellite repeats its orbit only every 46 days, so that repeat-pass interferometry can be implemented only with this minimum temporal baseline. However, such long baselines lead to large levels of temporal decorrelation, and hence are limiting for quantitative applications such as POLInSAR height retrieval and coherence tomography. Finally, we note the high orbit of the satellite, by which propagation effects through the ionosphere cannot be ignored.

To illustrate this we show data for calibration trihedral corner reflectors (see Figure 1.21) deployed in Adelaide, South Australia. The data was collected on 10 June 2006 at 13:50 UT. Figure 9.63 shows SAR images in HH (left) and HV (right) channels. This data has been calibrated using the scheme of Section 9.3.2, but has not been corrected for Faraday effects. From scattering theory we expect the HV backscatter to be zero for the trihedral, but we see it is only around 20 dB below the copolar signal. This is due to a small rotation of the plane of polarisation through the ionosphere. The level of Faraday rotation can be estimated using equation (4.83). The estimated level is around +3-degree one-way rotation. When this is removed from each pixel of the image we obtain the corrected imagery shown in Figure 9.64. Here we see much better isolation and residual cross-talk levels below -30 dB. This nicely illustrates how vector-wave propagation models can be used to improve data quality and system calibration for spaceborne radar systems.

One important application of such radars is in the identification and characterization of coherent (polarised, in our context) scattering points in radar imagery (Ferro-Famil, 2003; Schneider, 2006). Figure 9.65 is an example, in



Fig. 9.64 Satellite SAR images (HH left, and HV right) for a trihedral corner reflector after Faraday rotation correction



Fig. 9.65 HH image of ALOS-PALSAR corner reflector scene (Adelaide, South Australia)

which is shown an expanded view of the corner reflector scene. (The three corner reflectors are seen—that at far left being used in Figures 9.63 and 9.64.) We then use a 3×5 (range *x* azimuth) window centred on each pixel to estimate the local entropy from the eigenvalues of the coherency matrix. Figure 9.66 shows the distribution of entropy/alpha values for the whole scene. We note that the bulk of background pixels lie with high entropy, but there are several distinct bands of low entropy distributions with various alpha values. The trihedrals are seen lying close to the origin (small alpha and low entropy). However, there



Fig. 9.66 Distribution of entropy/alpha values for all pixels in Figure 9.69

are clearly coherent points (low entropy) with alpha values greater than 45°, indicating second-order dihedral scattering.

The differences in alpha then relate directly to the boundary conditions for that scattering process. In particular they relate primarily to the dielectric constant of the reflection. Furthermore, given the small angle of incidence of the data (typically a 24-degree angle of incidence in mid swath) and its small variation across the limited range swath, this dependence is primarily determined by one of the two surfaces involved, and not both (see Figure 3.15). For dry materials (low ε_r) we can even approach the Brewster angle for reflection on one of the surfaces, so that we can observe *either* 0- or 180-degree phase shift between HH and VV, depending on which side of the Brewster point we operate. A phase shift of 180° will be seen for dielectric constants greater than 5 (the Brewster angle of ALOS). Hence the alpha angle for dihedral scattering at small angles of incidence can be less than $\pi/4$. We can quantify this for the geometry of PALSAR by calculating the variation of the apparent alpha of the boundary conditions as a function of the dielectric constant.

Figure 9.67 shows an example. Here we show the variation of alpha with dielectric constant of the vertical surface (for $\varepsilon_r = 10$ for the horizontal). We show results for three angles of incidence—23°, 24°, and 25° (covering the swath of ALOS-PALSAR)—and see only a small change. We see that increasing alpha is directly linked to increasing dielectric constant of the vertical surface, and that we have high sensitivity to changes in dielectric constant. This provides us with a way to estimate ε_r of such dihedrals directly from spaceborne radar data. For example, we can use Figure 9.67 to relate alpha directly to dielectric constant using a suitable curve fit. One key feature of the fit must be that when α tends to $\pi/2$ so ε_r tends to infinity (a metallic dihedral has an alpha of $\pi/2$), and hence the function must have a pole at $\pi/2$. We then obtain the following

fit from Figure 9.67 (fitting $\theta = 24$ as the middle of the swath):

$$\varepsilon_r \approx \frac{3.2299}{\left(\frac{\pi}{2} - \alpha_s\right)^{2.5}} - 0.8522 \tag{9.46}$$

This relation then enables us to estimate the dielectric constant of the vertical surface reflection in dihedral scattering directly from the alpha angle of the dominant eigenvector.

These few examples help illustrate the potential of spaceborne radar polarimetry. The successful fusion of interferometry with polarimetry from space will have to await the development of single-pass spaceborne systems, or at least short temporal baseline low-frequency repeat pass sensors. At the time of writing, no such system is yet operational, but plans for the launch of Tandem-X in 2009 will see deployment of a spaceborne single-pass POLInSAR system at X-Band (Krieger, 2007). This will not only open new applications in the fusion of interferometry with polarimetry, but also provide new stimulus to continued study of our understanding and exploitation of the 'memory' effect imprinted on polarised electromagnetic waves scattered by random media.



Fig. 9.67 Relation between alpha and dielectric constant of vertical reflecting surface of dihedral for ALOS-PALSAR angles of incidence (23–25°)

Introduction to matrix algebra

A1

In this Appendix we gather together some basic definitions and concepts from matrix linear algebra. (For more details see Gershenfeld (1999), Strang (2004), and Press (2007).) We concentrate on those of particular importance to the subject matter of this book: namely, the matrix description of polarised wave scattering and propagation.

A1.1 Matrix definition, diagonal, upper and lower triangular forms

A *matrix* [A] (sometimes written in this book as a bold capital without brackets, A) is a rectangular array of numbers arranged with *m* rows and *n* columns. The *dimensions of the matrix* are then $m \ge n$. The general element located in row *i* and column *j* is termed a_{ij} . If all elements are zero expect when i = j, the matrix is termed *diagonal*. Two important variations of this idea are *upper and lower triangular* matrices—the former having zeros below the main diagonal, and the latter above. [A] can be a real or complex matrix, depending on whether one or more of the elements a_{ij} are real or complex numbers.

A1.2 Matrix multiplication, Kronecker sums and products

Two matrices can be multiplied only if they are size compatible; that is, two matrices [A] and [B] can be multiplied to generate $[C] = [A] \cdot [B]$ only when [A] is $m \times n$ and [B] is $n \times q$, so the product [C] will have dimensions $m \times q$. The elements of [C] c_{ij} are then formed from the inner (scalar) product of the *i*th row of [A] with the *j*th column of [B].

There are two other important ways in which matrix elements can be combined to generate a new matrix. The *Kronecker product* [*C*] is an *mp*-by-*nq* matrix derived from two matrices [*A*] which is $m \times n$ and [*B*] which is $p \times q$, as shown in equation (A1.1):

$$[C] = [A] \otimes [B] = \begin{bmatrix} a_{11}B & \cdots & a_{1n}B \\ \vdots & \ddots & \vdots \\ a_{m1}B & \cdots & a_{mn}B \end{bmatrix}$$
(A1.1)

As such, this operation replaces every element of [A] by its *ij* element multiplied by the whole matrix [B]. This form is particularly useful for converting

composite matrix products into a single matrix vector operation, so if we have a matrix [X] and two product matrices [A] and [B], as shown in equation (A1.2), then the vectorization of [X] is transformed by a Kronecker product matrix as shown:

$$[A][X][B] \equiv [A] \otimes [B]^T \underline{x} \tag{A1.2}$$

where \underline{x} is a lexicographic row vectorization of the elements of [X]:

$$[X] = \begin{bmatrix} x_{11} & x_{12} & \dots & x_{1n} \\ x_{21} & x_{22} & \dots & x_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n1} & x_{n2} & \dots & x_{nn} \end{bmatrix} \Rightarrow \underline{x} = \begin{bmatrix} x_{11} \\ x_{12} \\ \vdots \\ x_{nn} \end{bmatrix}$$
(A1.3)

The *Kronecker sum* of two square matrices $\mathbf{A}(n \ge n)$ and $\mathbf{B}(m \ge m)$ is likewise defined as shown in equation (A1.4):

$$[C] = [A] \oplus [B] = [A] \otimes [I_m] + [I_n] \otimes [B]$$
(A1.4)

where [I] is the m x m identity matrix, defined such that $[I] \cdot [A] = [A] \cdot [I] = [A]$.

A1.3 Square matrices, powers and the exponential

An important special class of matrices arises when m = n; that is, [A] has an equal number of rows and columns. In this case the matrix [A] is square, and can be multiplied with itself to generate power series; so, for example, $[A]^2 = [A] \cdot [A]$ is well defined, and also has dimension $m \times m$. One key consequence of this idea is that we can define matrix functions from their series representations. The most important example of this is the matrix exponential function, defined formally by the following series:

$$\exp([A]) = [I] + [A] + \frac{[A]^2}{2!} + \dots + \frac{[A]^n}{n!} + \dots$$
(A1.5)

This function is widely used in polarisation algebra (see Appendix 2 for more details). Finally, we note that for two different square matrices of the same dimension [A] and [B] (both $m \times m$), the products $[A] \cdot [B]$ and $[B] \cdot [A]$ are both well defined but are not necessarily equal. Hence matrix multiplication (unlike 'ordinary' multiplication) is non-commutative. The ability of matrices to represent differences in the order of multiplications is a key attraction for their use in problems involving rotations and transformations as encountered in polarimetry.

A1.4 Inverse matrix, minors and determinants

The *inverse* of a square matrix [A], denoted $[A]^{-1}$, is defined by the relationship $[A]^{-1} \cdot [A] = [A] \cdot [A]^{-1} = [I]$. Although this relationship applies for arbitrary

matrix dimension, special attention is paid to the case $m = 2-2 \times 2$ matrices for which the following simple expression can be used to directly calculate the inverse:

$$[A] = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \Rightarrow [A]^{-1} = \frac{1}{a_{11}a_{22} - a_{12}a_{21}} \begin{bmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{bmatrix}$$
(A1.6)

The multiplicative scale factor in front of the inverse matrix $a_{11}a_{22}-a_{12}a_{21}$ is called the *determinant* of the matrix $|\mathbf{A}|$ or det([A]), and evidently problems arise with the existence of the inverse when this determinant goes to zero. The concept of determinant is important, and can be extended to arbitrary matrix dimensions. In this case we define the determinant as a scalar obtained from an $n \times n$ matrix [A], as shown in equation (A1.7):

$$|A| = \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix} = \sum_{j=1}^{n} a_{ij} C_{ij} = \sum_{j=1}^{n} a_{ij} (-1)^{i+j} M_{ij}$$
(A1.7)

Here C_{ij} is called the *cofactor* associated with the element a_{ij} , and is in turn related to a signed version of M_{ij} , called the *minor*, associated with the element a_{ij} . The minor is itself a determinant, but importantly for a matrix of reduced dimension (n-1)—the matrix formed by eliminating the *i*th row and *j*th column of [A]. In this way we can reduce every determinant by reduction of dimension to calculation of a series of 2 x 2 sub-matrices. The summation on the righthand side of equation (A1.7) is called the Laplace expansion of determinant by the *i*th row (where *i* can be chosen arbitrarily). Similar expressions can be used for a Laplace expansion by the *j*th column, again involving minors and cofactors.

For arbitrary dimensions the inverse can be now be formally written as follows:

$$[A]^{-1} = \frac{adj([A])}{\det([A])} \quad adj([A]) = [C_{ij}]^T$$
(A1.8)

where adj([A] is called the *adjoint* matrix—itself formed as the transpose of the matrix of cofactors of [A]. The *transpose* of a matrix is an operation that exchanges rows and columns, formally replacing the *ij* element by the *ji* element. If a matrix [A] equals its transpose or $[A] = [A]^T$, then [A] is termed symmetric, having the property that $a_{ij} = a_{ji}$. Another special case is when $[A] = -[A]^T$, in which case the matrix is called *skew or anti-symmetric*. In this case it is easy to see that the diagonal elements of an antisymmetric matrix must all be zero.

Another important scalar function of matrix elements is the *trace* or Tr([A]). This is defined as the sum of the diagonal elements $Tr([A]) = \sum_{i=1}^{n} a_{ii}$. Two very important properties of trace used extensively in analytic studies follow from the ability to change the order of matrix products inside the trace operation, so that for dual products, for example, $Tr(\mathbf{AB}) = Tr(\mathbf{BA})$, and also the cyclic property of the trace for triple products so that Tr(ABC) = Tr(BCA) = Tr(CAB). One important consequence of this result is that the trace is invariant to unitary similarity transformations of a matrix [A], since $Tr(\mathbf{U}^{*T}\mathbf{A}\mathbf{U}) = Tr(\mathbf{A}\mathbf{U}\mathbf{U}^{*T}) = Tr(\mathbf{A})$. Hence the trace is equal to the sum of the eigenvalues of [A], and represents, in polarimetry, an important invariant quantity identified as the total scattered power by an object.

A1.5 Hermitian and anti-Hermitian matrices

For the case of complex matrices, another important operation can be defined by combining transpose with conjugation. Firstly we define the conjugate matrix $[A]^*$ as the matrix obtained by forming the complex conjugate of each element of [A]. By then combining this with the transpose we obtain the conjugate-transpose or *Hermitian adjoint* operation:

$$[B] = [A]^{*T} \Rightarrow b_{ij} = a_{ji}^* \tag{A1.9}$$

Note that sometimes in the literature the transpose symbol 'T' is omitted for notational convenience, and that only the conjugate sign is used to indicate both operations implicitly. There is then, of course, the possibility of ambiguity of notation. To counter this, the 'dagger' symbol \dagger is often used to represent the combined transpose and conjugate operations, so $[A]^{\dagger} = [A]^{*T}$. Again, an important class of matrices arises when $[A]^{\dagger} = [A]$. These are termed *Hermitian* matrices, and arise a great deal in the development of polarised wave scattering and propagation. Cleary, such matrices must have real elements along the diagonal and complex conjugate elements on matching off-diagonal locations. If $[A]^{\dagger} = -[A]$, then [A] is termed *skew or anti-Hermitian*, and again must have zeros along the diagonal.

A1.6 Orthogonal and unitary matrices

Two important classes of matrix can be defined by relating transpose and Hermitian adjoint operations to the matrix inverse. In the first case we define a matrix as orthogonal if $[A]^T \cdot [A] = [I]$ that is, if the inverse of the matrix is just its transpose. In this case the $n \times n$ matrix [A] can be decomposed into a set of *n* mutually orthogonal *n*-element column vectors <u>a</u>, where, for example, $\underline{a}_1^T \underline{a}_2 = 0$, and so on, as shown in equation (A1.10):

$$[A] = \begin{bmatrix} \underline{a}_1 & \underline{a}_2 & \dots & \underline{a}_n \end{bmatrix}$$
(A1.10)

In a related sense, if a matrix satisfies the relation $[A]^{*T}[A] = [I]$, then it is termed *unitary*, and again it can be decomposed into a set of n column vectors—this time with orthogonality defined in the complex Hermitian sense: $\underline{a}_1^{*T}\underline{a}_2 = 0$. Since orthogonality is a key physical concept in polarimetry, such matrices often arise in applications, particularly in the context of similarity transformations.

A1.7 Similarity and congruent transformations

Two $n \times n$ matrices [A] and [B] are called similar if there exists an invertible $n \times n$ matrix [P], so that [A] and [B] are related as follows:

$$[P]^{-1}[A][P] = [B]$$
(A1.11)

Another way to think about this relationship is that the matrix [A] is transformed into the matrix [B] by operation of [P]. This is termed a *similarity transformation* of [A] by [P]. Often the matrix [P] is unitary, in which case equation (A1.11) becomes a unitary similarity transformation and the matrices [A] and [B] are unitarily similar. Similar matrices share many properties in common. For example, the determinants are equal, so det([A]) = det([B]). More importantly, their eigenvalues are equal (although their eigenvectors are different). An important variation of this scheme is the congruent transformation. In this case, [A] and [B] are related again by a third matrix [P], but now in the form:

$$[P]^{T}[A][P] = [B]$$
(A1.12)

This is to be clearly distinguished from the more common similarity transformation of equation (A1.11), and arises in the description of the backscatter of polarised waves.

A1.8 Rayleigh's quotient, positive definite, positive semidefinite matrices

Another important class of composite products is the embedding of a matrix between two vectors to generate a scalar, as shown in equation (A1.13). Furthermore, if the matrix is real orthogonal [O], or complex Hermitian [H], then this scalar is always real. To see this, just take the transpose (for [O]) or conjugate transpose (for [H]), and note the invariance of the scalar to these operations, implying that the scalar must be real. When this scalar is greater than zero for all vectors \underline{x} , the matrix is termed *positive definite*, or PD. When the scalar is just non-negative (so including zero), the matrix is termed *positive semi-definite*, or PSD. These conditions are summarized in equation (A1.13):

$$\underline{x}^{T}[O]\underline{x} > 0 \qquad \underline{x}^{*T}[H]\underline{x} > 0 \qquad \text{PD}$$

$$\underline{x}^{T}[O]\underline{x} \ge 0 \qquad \underline{x}^{*T}[H]\underline{x} \ge 0 \qquad \text{PSD}$$
 (A1.13)

A classical problem in matrix algebra is to find the <u>x</u> vector that maximizes the scalar. To do this we need to first normalize by the magnitude of the vector to obtain Rayleigh's quotient R, as shown in equation (A1.14):

$$R\left(\underline{x}\right) = \frac{\underline{x}^{*T}[H]\underline{x}}{\underline{x}^{*T}\underline{x}}$$
(A1.14)

Such ratios arise a great deal in polarimetry applications. It is therefore of importance to be able to test if a matrix is PD or PSD, and to then to find a systematic way of maximising or minimising R. One of the best ways to do this is to use an eigenvalue approach, as we now show.

A1.9 Subspaces, eigenvectors and eigenvalues

We are often interested in a subset of the full set of vectors \underline{x} in relations such as equation (A1.14). One particularly important set form is the *nullspace* of the matrix [A]. These are by definition the set of vectors \underline{x}_N satisfying the following relation:

$$[A]\underline{x}_N = 0 \tag{A1.15}$$

A second important set are the *eigenvectors* \underline{e} . A square matrix [A] of dimension $n \times n$ has *n* such eigenvectors, defined by the following equation:

$$[A] \underline{e} = \lambda \underline{e}$$

$$\Rightarrow [A] \underline{e} - \lambda \underline{e} = 0 \qquad (A1.16)$$

$$\Rightarrow \det([A] - \lambda[I]) = 0$$

where λ is a scalar called the *eigenvalue*. In this sense we can consider [A] to operate on a vector x, and equation (A1.16) states that for some special vectors, e, this operation will leave the vector x unchanged, apart from multiplication by a scalar. This physical interpretation is useful in many applications in polarimetry, although equation (A1.16) is also an important general mathematical idea. The eigenvalues can be found from the zeros of the determinant (generally an *n*th order polynomial with *n* complex roots), as shown in equation (A1.16). It is of special interest to be able to express a general matrix [A] as a function of its eigenvectors and eigenvalues. An important theorem, called Schur's theorem, gives us a systematic way to achive this. Schur's theorem states that for any square matrix [A] there exists a unitary matrix [U], such that $[U]^{-1}[A][U] = [B]$ is upper triangular. The eigenvalues of [A] must be shared by the similar matrix [B] and appear along its main diagonal. The unitary matrix [U] is then composed of the eigenvectors of [A] as its columns. Hence we can write a general matrix [A] in terms of its eigenvalues and eigenvectors, as shown in equation (A1.17):

$$[A] = [U] \cdot \begin{bmatrix} \lambda_1 & \delta_{12} & \cdots & \delta_{1n} \\ 0 & \lambda_2 & \cdots & \delta_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix} \cdot [U]^{*T} \qquad [U] = \begin{bmatrix} \underline{e}_1 & \underline{e}_2 & \cdots & \underline{e}_n \end{bmatrix}$$
(A1.17)

An important class of matrices has an even simpler form, when the off-diagonal δ terms are all zero and the matrix [*B*] is diagonal. Symmetric and Hermitian matrices fall into this category. The proof for Hermitian matrices is very simple, since if [*A*] is Hermitian then so is [*B*], and the only way [*B*] can be Hermitian (equal to its conjugate transpose) is if it is diagonal, with real eigenvalues. Hence we have proved that Hermitian matrices have real eigenvalues and orthogonal eigenvectors—a result of great importance in polarimetry. In general, the required condition that [*B*] is diagonal is that the matrix [*A*] be normal, which by definition means that it commutes with its conjugate transpose; that is, $[A][A]^{*T} - [A]^{*T}[A] = 0$. Unitary matrices, for example, also satisfy

this condition (when the conjugate transpose equals the inverse), and so can also be expressed in diagonal form.

Eigenvalue decompositions are also of interest for solving optimization problems. Taking, for example, Rayleigh's quotient again (equation (A1.14)), we return to the problem of how to maximize this ratio. To solve this we can employ the *Lagrange multiplier* method of constrained optimization. We first set up a Lagrangian function L as follows:

$$L = \underline{x}^{*T}[H]\underline{x} - \lambda(\underline{x}^{*T}\underline{x} - 1)$$
(A1.18)

and then differentiate with respect to the variable \underline{x} to find the stationary points as zeros of the derivative, as shown in equation (A1.19):

$$\frac{dL}{d\underline{x}} = [H]\underline{x} - \lambda \underline{x} = 0 \tag{A1.19}$$

This implies that the optimum solution for \underline{x} is the eigenvector of [H] corresponding to the maximum eigenvalue λ_{max} . This simple example illustrates how eigenvalue decompositions and optimization theory can be formally linked.

Finally, we consider an extension of this concept to the *singular value decomposition*, or SVD. The Schur decomposition employs a single unitary matrix [U]. If, on the other hand, we employ two unitary matrices [U] and [V], then any matrix (even those that are not square) can be expressed as a function of a purely diagonal matrix [D], as follows:

$$[A] = [U] \cdot [D] \cdot [V]^{*T}$$

$$[D] = \begin{bmatrix} s_1 & & & \\ & \ddots & & \\ & & s_p & & \\ & & & 0 & \\ & & & \ddots & \\ & & & & 0 \end{bmatrix} |s_1| \ge |s_2| \cdots \ge |s_p| \quad (A1.20)$$

The diagonal elements of [D] are termed the *singular values* of the matrix [A] and [U], the left singular vectors, while [V] is the set of right singular vectors. In the general case when [A] has dimensions $m \times n$, [U] is $m \times m$, [D] is $m \times n$, and [V] is $n \times n$. However, only p elements of [D] are non-zero. The dimensions p and m-p define two important subspaces that have important applications, as we now consider.

A1.10 Norms, condition number, least squares and the SVD

Very often we wish to consider matrix equations of the following form:

$$[A] \underline{x} = \underline{b} \tag{A1.21}$$

There are two important subspaces to this problem. The *range* of the matrix [A] is the space of all possible vectors <u>b</u> for which the equation is solvable.

The dimension of the range is called the *rank*. Secondly, the set of vectors \underline{x} which satisfy the homogeneous equation ($\underline{b} = 0$) define the null space of [A]. If there is no null space, then [A] is of *full rank*. The SVD provides a useful perspective on these subspaces. The columns of [U] in equation (A1.20) that are associated with non-zero singular values form an orthonormal basis for the range of [A], while the columns of [V] associated with the zero singular values form an orthonormal basis for the null space.

These concepts are particularly important in the solution of least squares problems. In these cases we are usually interested in finding solution vector \underline{x} that solves an overdetermined set of equations. Given the possibility that an exact solution may not exist, we ask instead to find \underline{x} that has the minimum residual; that is, one that minimizes the function L, defined as:

$$L(\underline{x}) = \left| [A] \, \underline{x} - \underline{b} \right|^2 \tag{A1.22}$$

The classical way to solve this is to expand L and differentiate with respect to \underline{x} and set to zero. This yields the so-called *normal equations*, as follows:

$$\begin{aligned} A\underline{x} - \underline{b} \Big|^2 &= (A\underline{x} - \underline{b})^T (A\underline{x} - \underline{b}) \\ &= \underline{x}^T A^T A \underline{x} - 2\underline{x}^T A^T \underline{b} + \underline{b}^T \underline{b} \\ &\frac{\partial}{\partial \underline{x}} = 0 \Rightarrow A^T A \underline{x} - A^T \underline{b} = 0 \end{aligned}$$
(A1.23)
$$\Rightarrow \underline{x} = (A^T A)^{-1} A^T \underline{b}$$

There are two issues to be faced with such a solution. Firstly, does the inverse matrix exist at all; and secondly, how sensitive is the solution x to perturbations or small errors in the vector \underline{b} ? If there is any serious amplification of such errors we say the equations are *ill-conditioned*. The SVD provides an important insight into both these situations, as we now consider.

The generalized or pseudo-inverse of [A], designated $[A]^+$, can be defined from the SVD as follows. Writing the function L in terms of the SVD we obtain the following simplified expansion:

$$L = |UDV^{T}\underline{x} - \underline{b}|^{2} = |DV^{T}\underline{x} - U^{T}\underline{b}|^{2}$$

$$\frac{U^{T}\underline{b} = \underline{z}}{V^{T}\underline{x} = \underline{y}} \Rightarrow L = |D\underline{y} - \underline{z}| = \sum_{i=1}^{p} (s_{i}y_{i} - z_{i})^{2} + \sum_{i=p+1}^{m} z_{i}^{2}$$
(A1.24)

Minimising *L* is now seen to occur for the following choice:

$$y_i = \frac{z_i}{s_i}, \underline{x} = V \underline{y} \Rightarrow \underline{x} = V D^+ U^T \underline{b} = A^+ \underline{b}$$
(A1.25)

where D^+ is a rather strange kind of inverse matrix, formed as follows. When the diagonal elements of D are non-zero then we take the reciprocal, but when

they are zero we leave then as zero:

Equation (A1.25) then provides a formal solution to linear least squares problems, even in the case when the inverse of $\mathbf{A}^{T}\mathbf{A}$ in the normal equations does not exist. The generalized or pseudo-inverse of [*A*] is then defined as shown in equation (A1.25). When the singular values of [*A*] have a clear cut-off the above formulation is clear, but more often the situation is that the singular values fall off slowly and never actually go to zero. In this case we face another challenge. Here the matrix is technically of full rank, but practically we can obtain an ill-conditioned system. Again we can use the SVD to quantify this via the concept of condition number of the matrix [*A*] or CN(*A*) as follows.

We are interested in how any fractional errors in the input vector \underline{b} are mapped into errors in the solution vector \underline{x} . To obtain this we first define the *norm* of a matrix ||A||, which is a scalar defined in equation (A1.27):

$$\|A\| = \max_{\underline{x} \neq 0} \frac{|A\underline{x}|}{|\underline{x}|} \Rightarrow \|A\underline{x}\| \le \|A\| \|\underline{x}\|$$
(A1.27)

This norm has one very important property: the norm of a product is less than or equal to the product of the norms. We can find an expression for the norm of [A] by using the SVD and first generating the squared norm, as shown in equation (A1.28):

$$\|A\|^2 = \frac{\underline{x}^T A^T A \underline{x}}{\underline{x}^T \underline{x}} \tag{A1.28}$$

Now the matrix $A^{T}A$ is always symmetric, and so this Raleigh quotient is maximized by the maximum eigenvalue of $A^{T}A$ (see equation (A1.19)), λ_{max} . However, $A^{T}A = VD^{2}V^{T}$, and so this eigenvalue is equal to the squared modulus of the maximum singular value of [*A*]. Hence we finally have an expression for the norm of matrix [*A*] as shown in equation (A1.29):

$$\|A\| = |s_1| \tag{A1.29}$$

Returning to the issue of error amplification, we have $\underline{b} = A\underline{x}$, and so any error δx will satisfy $\delta \underline{x} = A^{-1}\delta b$. Using norms on both these equations, and re-ordering, we can then write an expression for the condition number *CN*, as shown in equation (A1.30):

$$\frac{\|\underline{b}\| \le \|A\| \cdot \|\underline{x}\|}{\|\delta\underline{x}\| \le \|A^{-1}\| \cdot \|\delta\underline{b}\|} \Biggr\} \Rightarrow \frac{\|\delta\underline{x}\|}{\|\underline{x}\|} \le CN \frac{\|\delta\underline{b}\|}{\|\underline{b}\|} \to CN = \|A^{-1}\| \cdot \|A\| = \left|\frac{s_{\max}}{s_{\min}}\right|$$
(A1.30)

Here we see that the amplification factor depends on the ratio of maximum to minimum singular values of [A]. It is easy to show that if [A] is unitary, for example, then CN = 1, and the errors are not amplified at all. In general, however, we find that CN > 1, and some care must be taken to ensure that errors are not amplified too much. Note from equation (A1.30) that we have a way of controlling this amplification by just setting any small singular values to zero in the pseudo-inverse, so increasing s_{\min} and reducing the CN until an acceptable ratio is achieved. The price to pay for this, however, is a rank reduction of [A].

Unitary and rotation groups

A2

In Appendix 1 we defined some very basic properties of matrices and matrix algebra as used in a description of polarised waves. Here we extend the discussion to examine several more advanced ideas, largely based on group theory, which are widely used in a description of the propagation and scattering of polarised electromagnetic waves, forming the subject of polarisation algebra.

There are three classical routes by which the transition from scalar to higher dimensional forms can be achieved mathematically (Murnaghan, 1962; Misner, 1973; Goldstein, 1980; Cornwell, 1984; Penrose, 1984; Rosen, 1995; Georgi, 1999). These are:

- Scalars_Vectors_Tensors
- Scalars_Complex Numbers_Quaternions_Bi-Quaternions
- Scalars_Spinors_Twistors

Each of these has its strengths and weaknesses in terms of ease of formulation, potential for quantitative analysis, and physical insight. However, two important general themes arise from all approaches: firstly, the relationship (sometimes conflict, as in relativistic quantum theory) between real and complex formulations of a problem (which in our context relate to the role of phase in multi-channel systems); and secondly, the unifying role played by group theory, which provides not only a convenient unifying framework, but also aids physical insight by exposing deep symmetries that can then be exploited to aid analysis of complicated problems.

Among the many concepts used in abstract algebra, one of the most useful is that of a group. We can introduce these ideas by identifying a hierarchy of algebraic concepts as follows, each one building on the properties of the simpler concepts to its left.

 $\text{SET} \Rightarrow \text{GROUP} \Rightarrow \text{FIELD} \Rightarrow \text{VECTOR SPACE} \Rightarrow \text{ALGEBRA}$

A group G is then defined as a set of elements together with a composition (generalized concept of a product) xy with $x, y \in G$, such that the following four conditions hold:

- a) Closure, $xy \in G$
- b) Associative, x(yz) = (xy)z
- c) Existence of the identity, xI = Ix = x
- d) Existence of the inverse, $xx^{-1} = I$

If in addition we have the property shown in e), then the group is called Abelian and such groups, as we shall see, have a simpler form and generate the basic building blocks of a general classification theory.

e) Commutative for group multiplication, xy = yx

We then specify G as finite or infinite if the number of elements is finite or infinite (called the order of the group). Again, finite and infinite groups have very different properties. Finally, if the elements of a group are functions of a continuous parameter—that is, rotation through an angle θ —then the group is called continuous.

It is remarkable that armed only with a set of such simple rules we can formalize many complex transformation problems and expose new and important underlying patterns in the description of polarised wave scattering.

Continuing the hierarchy, a *field* is a group with the extra concept of addition of elements, under which the field is commutative. There are three main fields of interest in physics and engineering: real numbers R, complex numbers C, and quaternions Q, all of which are important in polarisation algebra. An *algebra* itself is then a still more complicated structure, and consists of a group, a field, and three additional concepts: addition, scalar and vector multiplications. These, then, are the basic building blocks for polarisation algebra, as we now demonstrate.

We have seen that in the development of polarisation geometry, the mathematics of mapping from complex to real domains is of central significance. Two important examples are the SU(2)–SO(3) homomorphism, which underlies the geometry of the Poincaré sphere and the SL(2,C)–SO(3,1,R) homomorphism that leads to a Lorentz transformation of the Stokes vector and a real 4×4 matrix representation of scattering. In this Appendix we develop a general approach to parameterising complex unitary and real rotation groups of arbitrary dimension (Murnaghan, 1962; Cornwell, 1984; Cloude, 1995b; Rosen, 1995; Georgi, 1999). This formalism will clarify many of the features already discussed, and also highlight a third important mapping between SU(4) and SO(6), which can be used to provide a general physical interpretation of bistatic scattering in random media.

One of the most important practical examples of a continuous group is the general linear group formed from the set of $n \times n$ non-singular real GL(n,R) and complex GL(n,C) matrices. There is also a set of important sub-groups of GL such as SL(n), the special linear group of matrices with unit determinant; U(n,C) and SU(n,C), the set of unitary and special unitary complex matrices; and O(n,R)) and SO(n,R), the orthogonal and special orthogonal groups. A complete classification for simple continuous groups such as these was first developed independently by Sophus Lie in 1870 and Wilhelm Killing in 1880, and was refined by Elie Cartan in 1894 (Cartan, 1966). This classification leads to four infinite series of groups designated A, B, C, and D, together with five exceptional groups. These can be used to identify general mappings from complex into real domains, as we now show.

We begin with the concept of a Lie algebra L, named after the nineteenthcentury Norwegian mathematician Sophus Lie (1842–1899). This is an *n*dimensional linear vector space equipped with a Lie product or commutator defined between elements a and b, as shown in equation (A2.1):

$$[a,b] = ab - ba \tag{A2.1}$$

where group (matrix) multiplication is implicit in terms such as 'ab'. If [a,b] = 0 then the group is called Abelian. Sophus Lie was the first to show that the

properties of the algebra are embodied in a set of structure constants c_{pq}^{r} , defined from the commutation by

$$[a_{p}, a_{q}] = \sum_{r=1}^{n} c_{pq}^{r} a_{r}$$
(A2.2)

The connection between Lie algebras and groups is often provided by the matrix exponential function, so that we can define a general matrix element A as shown in equation (A2.3):

$$A = \exp(a) = I + a + \frac{a^2}{2!} + \frac{a^3}{3!} + \dots + \frac{a^n}{n!} + \dots$$
(A2.3)

Consider the following two important examples.

A2.1 The real Lie algebra L = su(n) and the group $SU(n)n \ge 2$

In this case **A** is unitary, and so *L* is the set of traceless anti-Hermitian $n \times n$ matrices, since

$$\exp(a)\exp(a)^{*T} = I$$
$$\det(\exp(a)) = 1$$
$$\Rightarrow \begin{cases} a^{*T} = -a\\Tr(a) = 0 \end{cases}$$
(A2.4)

The dimensionality of su(N) is $N^2 - 1$. For the algebra su(2) we then have the following three-dimensional representation with commutation relations as shown:

$$a_{1} = \frac{1}{2} \begin{bmatrix} 0 & i \\ i & 0 \end{bmatrix} \quad a_{2} = \frac{1}{2} \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \quad a_{3} = \frac{1}{2} \begin{bmatrix} i & 0 \\ 0 & -i \end{bmatrix}$$
(A2.5)
$$[a_{1}, a_{2}] = -a_{3} \quad [a_{2}, a_{3}] = -a_{1} \quad [a_{3}, a_{1}] = -a_{2}$$

By exponentiation we then arrive at the Pauli spin matrices $\sigma_i = -2ia_i$. The commutation properties of the Pauli matrices can be conveniently represented as a matrix, the *pq*th element of which is 1, 0 or -1 according to equation (A2.5), as shown in Table A.2.I.

Turning now to higher dimensions, the algebra su(3) is likewise constructed from eight basis matrices, as shown in equation (A2.6). Just as su(2) leads to the Pauli spin matrices as so-called *generators* for SU(2), so the corresponding set for the group SU(3) are the Hermitian Gell–Mann matrices, obtained as $\lambda_k = -ia_k$. Note that the scale factor in a_8 is used to ensure that for all products $Tr(a_ia_i) = -2\delta_{ij}$.

Finally, the algebra su(4) can be represented by the set of fifteen matrices shown in equation (A2.7). The corresponding generators for the group SU(4) are called Dirac matrices. The pattern is now clearly developed for representation of higher-dimensional unitary groups, although we see that the number of elements

Table A2.1 Commutation matrix of SU(2)				
σ	1	2	3	
1	0	1	-1	
2	-1	0	1	
3	1	-1	0	

quickly increases as we go to higher dimensions. We shall show later that there is an important way of classifying these groups based on a smaller dimensional set called the Cartan sub-algebra. However, we first consider a second important set of algebras related to the rotation groups.

A2.2 The real Lie algebra L = so(n) and the group SO(n)

In this case **A** is orthogonal, and so *L* is the set of traceless anti-symmetric $n \times n$ matrices, since

$$\exp(a) \exp(a)^T = I \\ \det(\exp(a)) = 1$$

$$\Rightarrow \begin{cases} a^T = -a \\ Tr(a) = 0 \end{cases}$$
 (A2.8)

Table A2.II shows a comparison of the dimensionality of this algebra compared to su(n). Also shown for completeness are the dimensions of other important classical matrix groups. An important example already encountered in polarimetry is so(3), which has a three-dimensional algebra formed from the following matrices:

$$a_{1} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{bmatrix} \quad a_{2} = \begin{bmatrix} 0 & 0 & -1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix} \quad a_{3} = \begin{bmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$
(A2.9)

Also shown in Table A2.II is the generalized Lorentz group SO(n,1,R), which again we have encountered in an interpretation of the geometry of the scattering matrix (see Section 1.5.3). This group combines *n*-dimensional rotations with a boost in one direction, and is formed from the set of real *n*+1 dimensional matrices satisfying the following equation involving the Lorentz metric:

$$[L] = \begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & -1 & \dots & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & \dots & -1 \end{bmatrix} = [M]^T [L] [M]$$
(A2.10)

For so(3,1) there is a homomorphism with sl(2,c), and the former can be represented by six matrices of the form shown in equation (A2.11):

Table A2.II	Dimensionality	of important	matrix groups
-------------	----------------	--------------	---------------

n	SO(n,R) 0.5n(n - 1)	SU(n,C) n ² - 1	$\frac{SL(n,C)}{2n(n^2-1)}$	SO(n,1,R) $0.5n(n^2+1)$
2	1	3	6	3
3	3	8	16	6
4	6	15	30	10
5	10	24	48	15
6	15	35	70	21

A2.3 The Killing form and Cartan matrix

To proceed towards a more general approach, we now consider construction of the Killing form for a Lie algebra L, named after Wilhelm Killing (1847–1923), and defined as shown in equation (A2.12):

$$B(x, y) = Tr \left\{ ad(x).ad(y) \right\}$$
(A2.12)

where *x* and *y* are elements of the algebra, and

$$ad(x) = \begin{bmatrix} x, x_j \end{bmatrix} \tag{A2.13}$$

is an $n \times n$ matrix, the *j*th row of which consists of the structure constants for the commutation of *x* with the *j*th element of *L*. The structure constants themselves form an *n*-dimensional representation of *L* called the adjoint representation. For example, for su(2) we have the following 3×3 matrix:

$$ad(a_{1}) = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{bmatrix} \quad ad(a_{2}) = \begin{bmatrix} 0 & 0 & -1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix} \quad ad(a_{3}) = \begin{bmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$
(A2.14)

The Killing form is then formed from the trace of matrix products, as shown in equation (A2.15):

$$B_{su(2)} = \begin{bmatrix} -2 & 0 & 0\\ 0 & -2 & 0\\ 0 & 0 & -2 \end{bmatrix} = -2\delta_{pq}$$
(A2.15)

The Killing form may initially seem a rather contrived concept, but it is used as the basic distinguishing feature of the Lie algebra, and is central to the general classification scheme. As mentioned earlier, it is possible to classify these algebras by identifying an important sub-algebra: the Cartan algebra H. This is Abelian by definition, and hence has very simple structure, being associated with a set of commuting or simultaneously diagonal matrices in the representation. Physically it can be considered a generalization of absolute phase in coherent signal analysis. When binary (two channel) operations are applied between signal channels this phase tends to disappear (in interferometry, for example), and we shall see that similar properties hold for the phase transformations generated by matrices obtained from the Cartan sub-algebra.

The dimension of *H* is called the rank of *L* and, importantly, this is always much smaller than the dimension of *L* itself. For su(n), for example, there are n-1 independent diagonal matrices, and hence the rank of su(n) is n-1. Hence su(2), su(3), and su(4) have rank 1, 2, and 3 respectively, while their dimensions (from Table A2.II) are 3, 8, and 15.

The algebra L can now be expressed as the direct sum of H, with rank k, and a remaining root subspace R, so that

$$L = H \oplus R \tag{A2.16}$$

Conveniently, the roots that span the space R can themselves always be written in terms of a subset of k simple roots r_s (where k equals the dimension of H), so that

$$r = \sum_{s=1}^{k} \alpha_s r_s \tag{A2.17}$$

where the coefficients α_s are generally complex. The classification can then be generated by constructing an $r \times r$ matrix from the simple roots, called the Cartan matrix A. The *jk*th element of A is obtained from the simple roots a_j and a_k as

$$A_{jk} = \frac{2B\left(a_j, a_k\right)}{B\left(a_j, a_j\right)} \tag{A2.18}$$

where B(..) is the Killing form. Clearly, the diagonal elements of A are equal to 2, but less obvious is that the off-diagonal elements are limited in value to 0, -1, -2, or -3. The Cartan matrix can be constructed from knowledge of the roots and *vice versa*; that is, we can construct the whole algebra from knowledge of A. Hence this matrix is the 'signature' of the algebra, and allows us to compactly describe higher dimensional groups.

Two examples will illustrate the method. We start with su(2), which has basis a_1, a_2 , and a_3 (equation (A2.5)), and the Cartan sub-algebra, which is onedimensional, with $h_1 = a_3$. From equation (A2.5) we then have the following (quasi-)eigenvalue conditions for the non-zero roots:

$$[h_1, (a_1 + ia_2)] = i (a_1 + ia_2)$$

$$[h_1, (a_1 - ia_2)] = -i (a_1 - ia_2)$$
(A2.19)

There are consequently two roots α_1 and $-\alpha_1$ with $\alpha(h_1) = i$. The onedimensional root subspaces are then defined from complex linear combinations of a_1 and a_2 as $\lambda(a_1 + ia_2)$ and $\mu(a_1 - ia_2)$, where λ and μ are arbitrary complex numbers. The Cartan matrix for su(2) has only one element: A = 2.

Turning to the more complicated case of su(3), from equation (A2.6) the Killing form is now

$$B(a_p, a_q) = -12\delta_{pq} \quad p, q = 1, 2, \dots 8 \tag{A2.20}$$

and with rank 2, the Cartan sub-algebra is spanned by $h_1 = a_3$ and $h_2 = a_8$ (see equation (A2.6)). It follows that we can write the following equations:

$$[h_1, a_2 - ia_1] = 2(a_2 - ia_1) \quad [h_2, a_2 - ia_1] = 0$$

$$[h_1, a_7 - ia_6] = -1(a_7 - ia_6) \quad [h_2, a_7 - ia_6] = \sqrt{3} (a_7 - ia_6)$$

$$[h_1, a_5 - ia_4] = 1(a_5 - ia_4) \quad [h_2, a_5 - ia_4] = \sqrt{3} (a_5 - ia_4)$$

$$[h_1, -a_2 - ia_1] = -2(-a_2 - ia_1) \quad [h_2, -a_2 - ia_1] = 0$$

$$[h_1, -a_7 - ia_6] = 1(-a_7 - ia_6) \quad [h_2, -a_7 - ia_6] = -\sqrt{3}(-a_7 - ia_6)$$

$$[h_1, -a_5 - ia_4] = -1(-a_5 - ia_4) \quad [h_2, -a_5 - ia_4] = -\sqrt{3}(-a_5 - ia_4)$$

$$(A2.21)$$

There are therefore six roots, $\alpha_1, \alpha_2, \alpha_3, -\alpha_1, \alpha_2$, and $-\alpha_3$, which can all be expressed as linear combinations of two simple roots: $r_1 = ah_1 + bh_2$. The coefficients *a* and *b* can be obtained from the commutation properties of the roots combined with the Killing form to generate a pair of simultaneous equations of the following form:

$$aB(h_1, h_j) + bB(h_2, h_j) = \alpha(h_j) \quad j = 1, 2$$

$$\alpha_1(h_1) = 2 \quad \alpha_1(h_2) = 0$$

$$\alpha_2(h_1) = -1 \quad \alpha_2(h_2) = \sqrt{3}$$

$$\alpha_3(h_1) = 1 \quad \alpha_3(h_2) = \sqrt{3}$$

(A2.22)

The coefficients a and b can then be calculated as

$$\alpha_{1} = \frac{1}{6}h_{1} = \frac{1}{6}\begin{bmatrix} 1 & 0 & 0\\ 0 & -1 & 0\\ 0 & 0 & 0 \end{bmatrix}$$

$$\alpha_{2} = -\frac{1}{12}h_{1} + \frac{\sqrt{3}}{12}h_{2} = \frac{1}{6}\begin{bmatrix} 0 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & -1 \end{bmatrix}$$
(A2.23)
$$\alpha_{3} = \frac{1}{12}h_{1} + \frac{\sqrt{3}}{12}h_{2} = \frac{1}{6}\begin{bmatrix} 1 & 0 & 0\\ 0 & 0 & 0\\ 0 & 0 & -1 \end{bmatrix}$$

where we have used the fact that $B(h_1, h_1) = B(h_2, h_2) = 12$ and $B(h_1, h_2) = 0$. From these we can calculate the following Killing forms and Cartan matrix:

$$B(\alpha_1, \alpha_1) = \frac{1}{3} B(\alpha_1, \alpha_2) = -\frac{1}{6} B(\alpha_1, \alpha_2) = -\frac{1}{6} B(\alpha_2, \alpha_2) = \frac{1}{3}$$

$$A = \begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix}$$
(A2.24)

One important point is that we can now reverse this procedure and use the Cartan matrix to generate the roots and hence the whole algebra. This is facilitated through use of a geometrical construction called Dynkin diagrams. These will finally lead us to the complete classification scheme.

A2.4 Dynkin diagrams: classification of unitary and rotation groups

The construction of Dynkin diagrams provides a convenient geometrical method for classifying Lie algebras (Cornwell, 1984; Georgi, 1999). This procedure is only part of a more general geometrical approach to the study of Lie algebras that involves the association of roots with vectors in a Euclidean space, with the Killing form employed as a scalar product. The commutation properties of the roots, together with the restrictive integer range for the Killing

form, mean that the set of non-zero vectors in this space is very restricted. In this way we can construct the full set of rank N spaces by employing purely geometrical methods.

A Dynkin diagram is constructed for each algebra *L* by associating a node with each simple root and connecting nodes corresponding to roots a_j and a_k by a number of lines given by $A_{jk}A_{kj}$, where A_{jk} is the *jk*th element of the Cartan matrix. Each node is also given a weight $\omega_j = \omega \langle \alpha_j, \alpha_j \rangle$, where ω is a constant chosen such that the minimum weight is unity. These diagrams are used to generate the whole classification scheme by starting with a root space of rank 1 and using an iterative scheme to generate root spaces of higher rank.

This scheme leads to classification of the four infinite sets of algebras, denoted A_i , B_i , C_i , and D_i with Dynkin diagrams shown in Figure A2.1. There are also five exceptional algebras E_6 , E_7 , E_8 , F_4 , and G_2 , with irregular Dynkin diagrams as shown in Figure A2.2. The classical continuous groups associated with the four infinite sets of Lie algebras through the matrix exponential function







 E_6

 E_7

Fig. A2.2 Dynkin diagrams for the 5 exceptional algebras
can then be identified as follows:

$$A_{N-1} \Rightarrow SU(N)$$

$$B_N \Rightarrow SO(2N+1)$$

$$C_N \Rightarrow USp(2N)$$

$$D_N \Rightarrow SO(2N)$$
(A2.25)

where USp(N) is a unitary group but with a symplectic inner product in N dimensions. (These symplectic groups are related to quaternions, and involve skew-symmetric bilinear forms; see Cornwell (1984) for more details.) The key concept is that for a homomorphism to exist between the various groups they must have the same Dynkin diagram, and their algebras are then isomorphic. In this way the constructs in Figure A2.1 can be used to identify higher-dimensional mappings from complex to real groups. We have seen in this book that such mappings are central to polarisation algebra.

Consider the following important examples. If $L = A_1$, the Dynkin diagram is a single node with unit weight. In this case we can construct the Cartan matrix as A = 2, and assign a one-dimensional root space. However, we also note that B_1 and C_1 have the same Dynkin diagram. Hence the three algebras are isomorphic. This result leads to the SU(2)–SO(3) homomorphism underpinning the geometry of the Poincaré sphere. Extending this, if $L = A_3$ then the Cartan matrix is of the following form:

$$A = \begin{bmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{bmatrix}$$
(A2.26)

which is the same as that for D_3 . Hence there is also a homomorphism between the groups associated with A_3 and D_3 . Finally, we see that $B_2 = C_2$, and with this we have exhausted all possible isomorphisms between the algebras. These important results are summarized in Table A2.III. The SU(2)–SO(3) and SU(4)–SO(6) homomorphisms are of particular interest in polarimetry studies. The SU(2) example is well known from studies of the Poincaré sphere, and so here we summarize the main details of the less well known SU(4) mapping; that is, given an element U_4 of SU(4,C) generate an equivalent element O_6 of SO(6). The algebra su(4) has rank 3, and a suitable basis for the Cartan sub-algebra can be obtained from equation (A2.7):

$$h_{1} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix} \quad h_{2} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix} \quad h_{3} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$
(A2.27)

Table A2.III Important homomorphic (isomorphic)relationships between Lie groups (algebras)

Algebras	Group Homomorphism	Dimension
A1 = B1	SU(2)-SO(3)	3
B2 = C2	Sp(2)-SO(5)	10
A3 = D3	SU(4)-SO(6)	15

The Dynkin diagram and corresponding Cartan matrix are then of the form shown in equation (A2.28):

$$\begin{array}{cccc} 1 & 1 & 1 \\ \bigcirc & & & \\ \alpha_1 & \alpha_2 & \alpha_3 \end{array} \qquad A = \begin{bmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{bmatrix}$$
 (A2.28)

A suitable set of generators for the group SU(4) are given by $\eta_k = -ia_k$, where a_k are defined in equation (A2.7). These fifteen matrices have the following commutation properties:

$$[\eta_a, \eta_b] = \sum_{c=1}^{15} 2i\varepsilon_{abc}\eta_c \tag{A2.29}$$

where the permutation symbol may be represented by a 15 × 15 matrix, as shown in Table A2.IV. The *ij*th element of this matrix is zero if η_i and η_j commute, and ±1 depending on the sign of the non-commuting elements. For SU(2) the corresponding matrix is shown in Table A2.I. To illustrate the power of this theory we consider the detailed mapping from SU(4) to SO(6), which is performed in three distinct stages, as follows.

SU(4)–SO(6) homomorphism stage 1

We begin by considering two vector spaces U and V. The tensor product $U \otimes V$ consists of a new vector space with basis $u_i \otimes v_j$ where i = 1, 2, 3 ... N, j = 1, 2, 3 ... M, and N and M are the dimensions (dim) of U and V. Consequently, dim $(U \otimes V) = \dim(U) \dim(V) = MN$. Typically we take repeated *r*th-order tensor products of a space with itself: $U \otimes U \otimes U ... = L^r(U_N)$. An element in this space is known as an *r*th-order tensor, and $L^r(U_N)$ is called the carrier space for an *r*th-order tensor.

Special consideration is given to second-order tensor space $L^2(U_N)$, were it is possible to form new tensors as linear combinations of the basis vectors $\underline{e}_i, \underline{e}_j$, which are antisymmetric under an interchange of subscripts. We do this

Table A2.IV	Commutation	matrix for	SU(4)
-------------	-------------	------------	-------

η	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1	0	-1	1	0	0	-1	1	0	0	-1	1	0	0	-1	1
2	1	0	-1	0	1	0	-1	0	1	0	-1	0	1	0	-1
3	-1	1	0	0	-1	1	0	0	-1	1	0	0	-1	1	0
4	0	0	0	0	0	0	0	1	1	1	1	-1	-1	-1	-1
5	0	-1	1	0	0	-1	1	1	1	0	0	-1	-1	0	0
6	1	0	-1	0	1	0	-1	1	0	1	0	-1	0	-1	0
7	-1	1	0	0	-1	1	0	1	0	0	1	-1	0	0	-1
8	0	0	0	-1	-1	-1	-1	0	0	0	0	1	1	1	1
9	0	-1	1	-1	-1	0	0	0	0	-1	1	1	-1	0	0
10	1	0	-1	-1	0	-1	0	0	1	0	-1	1	0	1	0
11	-1	1	0	-1	0	0	-1	0	-1	1	0	1	0	0	1
12	0	0	0	1	1	1	1	-1	-1	-1	-1	0	0	0	0
13	0	-1	1	1	1	0	0	-1	1	0	0	0	0	-1	1
14	1	0	-1	1	0	1	0	-1	0	-1	0	0	1	0	-1
15	-1	1	0	1	0	0	1	-1	0	0	-1	0	-1	1	0

by forming a wedge product or bivector defined as $\underline{e}_i \wedge \underline{e}_j = \underline{e}_i \otimes \underline{e}_j - \underline{e}_j \otimes \underline{e}_i = -\underline{e}_j \wedge \underline{e}_i$. Note that this implies that $\underline{e}_i \wedge \underline{e}_i = 0$. In general, the number of basis vectors existing in a fully antisymmetric subspace of $L^r(U_N)$ is N!/(r!(N–r)!). For a second-order space there are therefore N(N–1)/2 basis vectors. Significantly for us, if we start with U = C4–that is, N = 4, a four-dimensional complex space—then this results in six dimensions. This allows us the first part of our objective by mapping from C4 to C6 as follows.

From the basis vectors in C4, $\underline{e}_1, \underline{e}_2, \underline{e}_3$, and \underline{e}_4 , we first generate a sixdimensional space with basis vectors t_i generated from all possible wedge products given by equation (A2.30):

$$t_1 = e_1 \wedge e_2 \quad t_2 = e_2 \wedge e_3 \quad t_3 = e_3 \wedge e_1 t_4 = e_3 \wedge e_4 \quad t_5 = e_1 \wedge e_4 \quad t_6 = e_2 \wedge e_4$$
(A2.30)

SU(4)–SO(6) homomorphism stage 2

We now consider a general 4×4 complex matrix A in C4, and note how it maps into this new six-dimensional space C6. This can be obtained explicitly as shown in equation (A2.31):

$$\underline{e}_i \wedge \underline{e}_j \Rightarrow [A] \, \underline{e}_i \wedge [A] \, \underline{e}_j = \sum a_{ri} \underline{e}_i \wedge \sum a_{sj} \underline{e}_j = \sum \sum a_{ri} a_{sj} \underline{e}_r \wedge \underline{e}_s$$
(A2.31)

This mapping corresponds to a 6×6 matrix W, the 36 elements of which are derived from the 362×2 minors of A. In general, the *ij*th element of W is then the minor formed from the *i*th and *j*th rows with the *i*th and *j*th columns, so, for example, $w_{11} = a_{11}a_{22} - a_{12}a_{21}$, and so on.

Given any 4×4 unitary matrix U_4 , therefore, we can generate a 6×6 unitary matrix by calculating the 36.2×2 determinants of U_4 ; that is, the 1,1 element of U_6 is the 2×2 determinant from columns 1,2 and rows 1,2 of U_4 . Similarly, the 1,2 element is formed from the determinant from columns 1,2 and rows 2,3, and so on. In this way we achieve a mapping from C4 into C6.

SU(4)–SO(6) homomorphism stage 3

In general, $r \times r$ minors are involved in the exterior algebra $L^r(U_N)$. A special case arises when r = N, when there is only one basis vector, called the volume element associated with the basis set $\underline{e}_1, \underline{e}_2 \dots \underline{e}_N$. In our case we have, for $L^4(U_4)$, the following result;

$$[A] \underline{e}_1 \wedge [A] \underline{e}_2 \wedge [A] \underline{e}_3 \wedge [A] \underline{e}_4 = \det([A]) \underline{e}_1 \wedge \underline{e}_2 \wedge \underline{e}_3 \wedge \underline{e}_4 \qquad (A2.32)$$

Hence when $[A] \in SU(4)$ it has unit determinant and therefore when all four terms in the volume element are distinct their coefficient is 1, whereas the coefficient is 0 for all combinations with repeated indices.

This result is important, because we can now consider generation of a scalar, the inner product of two bivectors \underline{x} and y, as shown in equation (A2.33):

$$\frac{x}{y} = \sum x_i t_i \\ \underline{y} = \sum y_j t_j \end{cases} \Rightarrow \underline{x} \land \underline{y} = f(x, y) \underline{e}_1 \land \underline{e}_2 \land \underline{e}_3 \land \underline{e}_4$$

$$f(x, y) = x_1 y_4 + x_2 y_5 + x_3 y_6 + x_4 y_1 + x_5 y_2 + x_6 y_3$$
(A2.33)

This can be shown quite easily by explicit expansion using the basis vectors defined in equation (A2.30). Note that there is a cyclic permutation of ordering in y, and so to obtain a scalar product in C6 we consider not matrix products such as $\overline{W}^T W$ but must also include a permutation matrix P, as shown in equation (A2.34), where I_3 is the 3 × 3 identity matrix:

$$[P] = \begin{bmatrix} 0 & I_3 \\ I_3 & 0 \end{bmatrix} \Rightarrow [W]^T [P][W]$$
(A2.34)

Since we are considering SU(4), the matrix product will then have 1 in positions 14,25 and so on, where four distinct basis vectors occur and 0 elsewhere. In other words, the following matrix identity holds:

$$[W]^T [P][W] = [P]$$
(A2.35)

This result introduces the permutation matrix P, which has the useful property that $P^2 = I_6$ and is central to the next and final stage of our mapping procedure. Recall that so far we can go from a 4×4 unitary matrix A to a 6×6 matrix W, which is also unitary. (To prove this, use the fact that A maps to W, A^{*T} maps to W^{*T} , and then finally, I_4 maps to I_6 , to show that $W^{*T}W = I_6$.) However, we seek a mapping from SU(4) to SO(6), which we know exists from the associated Dynkin diagrams. We now therefore seek a similarity transformation to convert W into a real orthogonal matrix; that is, we seek a 6×6 matrix Q such that $R = QWQ^{-1}$ is real orthogonal. We can construct such a matrix using P, as follows.

We start by defining a symmetric matrix $\boldsymbol{Q} = \boldsymbol{Q}^T$, as shown in equation (A2.36):

$$[Q]^2 = [P] \Rightarrow [Q]^4 = [P]^2 = [I_6] \Rightarrow [Q]^{-1} = [Q]^3$$
 (A2.36)

The reason for this choice becomes clear when we consider testing whether **R** is orthogonal, as shown in equation (A2.37). We see that by using the properties in equation (A2.36) we guarantee that **R** is orthogonal, as required.

$$[R]^{T}[R] = \left(QWQ^{-1}\right)^{T} \left(QWQ^{-1}\right) = Q^{3}W^{T}PWQ^{3} = Q^{3}PQ^{3} = I_{6}$$
(A2.37)

Technically, \mathbf{R} could still be complex orthogonal, but we can show that it is real orthogonal by demonstrating that it is unitary as well as orthogonal; that is, that $\mathbf{R}^{*T}\mathbf{R} = \mathbf{I}_6$. This follows from a similar expansion to that shown in equation (A2.37). Hence **R** is real orthogonal as required.

The matrix \mathbf{Q} can then be defined explicitly, as shown in equation (A2.38):

$$Q = \frac{(1-i)}{2} \begin{bmatrix} I_3 & iI_3 \\ iI_3 & I_3 \end{bmatrix} \Rightarrow Q^2 = -i \begin{bmatrix} I_3 & iI_3 \\ iI_3 & I_3 \end{bmatrix} \cdot \begin{bmatrix} I_3 & iI_3 \\ iI_3 & I_3 \end{bmatrix} = \begin{bmatrix} 0 & I_3 \\ I_3 & 0 \end{bmatrix} = P$$
(A2.38)

This finally brings us to a general algorithm for mapping elements of SU(4) into corresponding elements of SO(6). We start by generating the 6×6 unitary

matrix $W = U_6$ from stage 2. If U_6 is then partitioned as shown in equation (A2.39),

$$U_6 = \begin{bmatrix} A & B \\ C & D \end{bmatrix}$$
(A2.39)

where A, B, C, and D are 3×3 sub-matrices, then we can always generate a 6×6 real orthogonal matrix by the following transformation:

$$Q = \frac{1-i}{2} \begin{bmatrix} I_3 & iI_3 \\ iI_3 & I_3 \end{bmatrix}$$

$$\Rightarrow O_6 = QU_6Q^{-1} = \frac{i}{2} \begin{bmatrix} B - C + i(A+D) & A - D + i(B+C) \\ D - A + i(B+C) & C - B + i(A+D) \end{bmatrix}$$
(A2.40)

For example, consider the element of U_4 given by a simple phase shift between elements in C4, as shown on the left of equation (A2.41). This maps into an 'equivalent' 6×6 real orthogonal matrix O_6 , as shown. The generator for this U_4 matrix is $-ia_{15}$ (see equation (A2.7)), and it maps into a rotation in the 2,5 plane. This is shown in equation (A2.7) as a subscript of the form $[\ldots]_{25}$.

Equation (A2.42) shows a second example: mapping a (real) 1,6 plane rotation into a (complex) unitary matrix with generator $-ia_6$. This procedure can be extended to all fifteen of the generators of SU(4). The plane rotations corresponding to each of the fifteen elements of su(4) are shown as subscripts in equation (A2.7).

$$U_{4} = \begin{bmatrix} e^{i\phi} & 0 & 0 & 0\\ 0 & e^{-i\phi} & 0 & 0\\ 0 & 0 & e^{-i\phi} & 0\\ 0 & 0 & 0 & e^{i\phi} \end{bmatrix} \Rightarrow O_{6} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0\\ 0 & \cos 2\phi & 0 & 0 & -\sin 2\phi & 0\\ 0 & 0 & 1 & 0 & 0 & 0\\ 0 & \sin 2\phi & 0 & 0 & \cos 2\phi & 0\\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$
(A2.41)

$$O_{6} = \begin{bmatrix} \cos\phi & 0 & 0 & 0 & -\sin\phi \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ \sin\phi & 0 & 0 & 0 & \cos\phi \end{bmatrix}$$
$$\Rightarrow U_{4} = \begin{bmatrix} \cos\frac{\phi}{2} & 0 & 0 & \sin\frac{\phi}{2} \\ 0 & \cos\frac{\phi}{2} & i\sin\frac{\phi}{2} & 0 \\ 0 & i\sin\frac{\phi}{2} & \cos\frac{\phi}{2} & 0 \\ -\sin\frac{\phi}{2} & 0 & 0 & \cos\frac{\phi}{2} \end{bmatrix}$$
(A2.42)

Coherent stochastic signal analysis

A3

In any discussion of noise processes, prime consideration is usually given to Gaussian random variables. Their importance stems from the Central Limit Theorem, which briefly states that the sum of a large number of independent and identically distributed random variables will be normally distributed. Due to the generic nature of this theorem, Gaussian statistics are often encountered in random wave propagation and scattering problems. (Some aspects of non-Gaussian statistics have been explored for polarised waves; see, for example, Bates (1998).) In this Appendix we briefly summarize the main impact of such stochastic models on coherent signal analysis, particularly on phase and coherence statistics (Touzi, 1999; Lee, 1994b, 2008; Lopez-Martinez, 2005; Ferro-Famil, 2008).

Signals generated by Gaussian random processes are characterized by statistical and not deterministic measures. Consequently, any single sample of such a process essentially contains zero information, and it is only by obtaining multiple samples and forming sums or integrals that the signal can be characterized by its statistical moments. The value of x for any particular sample is then independent of any previous values, and is taken from a normal distribution such that it is characterized by a probability density function (pdf) p(x), as summarized in equation (A3.1):

$$p(x) = G(m,\sigma) = \frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{(x-m)^2}{2\sigma^2}} \Rightarrow \begin{cases} \overline{x} = E(x) = \int_{-\infty}^{\infty} x.p(x) \, dx = m\\ \overline{x^2} = \int_{-\infty}^{\infty} x^2.p(x) \, dx = \sigma^2 + m^2\\ \sigma_x^2 = E((x-m)^2) = \sigma^2 \end{cases}$$
(A3.1)

For Gaussian signals the process is fully characterized by the mean *m* and standard deviation σ . Pure noise signals are often characterized by a zero mean process, written in shorthand as $G(0,\sigma)$. Hence Gaussian noise has only one free parameter, σ . Figure A3.1 shows an example of a signal composed of 256 samples taken from a G(0,1) random number generator. Such generators provide the basis for modelling multi-channel polarimetric and interferometric signals, and are a useful way to investigate the statistical properties of wave depolarisation, as we show later in this Appendix. First, however, we need to extend the simple scalar model of equation (A3.1) to complex signals, and then to the multidimensional complex signal vectors encountered in polarimetry.



Fig. A3.1 256 noise signal samples generated from a G(0,1) process

For complex signals, with amplitude and phase, as encountered in wave propagation and scattering, we must extend these ideas to account for both real and imaginary components of the signal. This is summarized in equation (A3.2). One of the most important relations in equation (A3.2) is that the expectation of the product *between* real and imaginary parts is zero. This is just a consequence of the fact that noise carries zero phase information, and hence the real and imaginary parts are independent Gaussian random variables. The phase, therefore, has a uniform distribution in the range 0 to 2π , while the intensity, defined as shown in equation (A3.2), has an exponential distribution. The amplitude *A* or square root of the intensity has a Rayleigh distribution as shown. Both have large variances, and so some care is required to minimize errors when estimating parameters from real data.

As a measure of the fluctuations in such data, the coefficient of variation (CV) can be defined as shown in equation (A3.3):

$$CV = \frac{1}{ENL} = \frac{\text{standard deviation}}{\text{mean}}$$
 (A3.3)

This coefficient is also related—as shown in equation (A3.3)—to the 'effective number of looks', or ENL, which is widely used in SAR image analysis. We see, for example, that for exponentially distributed data as in equation (A3.2), the ratio CV = 1. This emphasises that these fluctuations are not due to thermal noise. As the signal strength increases, so its variance also increases to keep CV = 1. Such fluctuations therefore cannot be reduced by increasing signal power. These fluctuations are common to all types of coherent imaging, where they are termed speckle noise (Lee, 1994a, 2008).

The simplest way to reduce speckle—the variance of the estimate, and therefore reduce the CV (increase ENL)—is to employ multi-look averaging. Here Lindependent samples are summed to obtain an estimate of the mean intensity. In this case the intensity distribution has a chi-square distribution with 2L degrees of freedom, as shown in equation (A3.4):

$$p(I) = \frac{L^{L}I^{L-1}}{(L-1)!\sigma^{2L}}e^{\frac{-LI}{\sigma^{2}}} \quad I \ge 0 \Rightarrow \begin{cases} E(I) = \sigma^{2} \\ \operatorname{var}(I) = \frac{\sigma^{4}}{L} \end{cases}$$
(A3.4)

The ratio of the standard deviation to the mean is then reduced to $1/\sqrt{L}$. This observation forms the basis for the design of speckle filters in SAR imaging. For example, the Lee filter (Lee, 1994a) proposes that we estimate the intensity of a pixel \hat{I} using a mixture of the pixel value itself, I, and its local mean, \overline{I} (usually estimated locally using a small M × N window centred on the pixel), based on a local Taylor expansion for the intensity of a pixel of the form shown in equation (A3.5):

$$\widehat{I} = \overline{I} + k(I - \overline{I}) \Rightarrow k = \frac{CV^2 - \frac{1}{L}}{CV^2}$$
(A3.5)

If the area is homogeneous (called fully developed speckle in the SAR context) then from equation (A3.5), k = 0, and the mean value is taken. However, if the region is very heterogeneous (a point target, for example) then *CV* will be much greater than 1, and so k = 1, and the local value is kept. In this way, local statistics can be used to strike a balance between spatial and radiometric resolution. Note that application of this approach to image data relies on two key assumptions: ergodicity in the mean, and local wide-sense stationarity—that space and time averages converge to the same mean value so that the spatial averaging locally around a SAR pixel can be considered equivalent to obtaining multiple samples of the same random process.

We have seen in Chapter 1 that polarimetry involves a two-dimensional complex space C2. Hence we need to take one further step in our characterization of noise by considering the statistical properties of signals in C2: pairs of complex signals of the form shown in equation (A3.6):

$$s_{1} = s_{1I} + is_{1Q} \in G_{c}(0, \sigma)$$

$$s_{2} = s_{2I} + is_{2Q} \in G_{c}(0, \sigma)$$
(A3.6)

We have seen, for example, that the product $s_1 \cdot s_2^*$ arises in many applications. This product is important, because the conjugate sign implies the phase *difference* between signals 1 and 2. Hence, while s_1 and s_2 individually have random phase, the phase difference can still be deterministic. *Noise* processes in C2 must therefore be further specified by the following added constraints:

$$E(s_1s_1^*) = \sigma^2 \quad E(s_1s_2^*) = 0 \quad E(s_2s_2^*) = \sigma^2$$
 (A3.7)

where again the zero expectation of the cross terms forces the phase difference to be uniformly random. We can summarize these properties of noise signals in C2 by generating a 2×2 covariance matrix from the expectation of the outer product of a vector in C2 with its conjugate, as shown in equation (A3.8):

$$[C]_{noise} = E\left(\begin{bmatrix} s_1\\ s_2 \end{bmatrix}, \begin{bmatrix} s_1 & s_2 \end{bmatrix}^*\right)$$
$$= \begin{bmatrix} E(s_1s_1^*) & E(s_1s_2^*)\\ E(s_2s_1^*) & E(s_2s_2^*) \end{bmatrix} = \begin{bmatrix} \sigma^2 & 0\\ 0 & \sigma^2 \end{bmatrix} = \sigma^2 \begin{bmatrix} 1 & 0\\ 0 & 1 \end{bmatrix}$$
(A3.8)

To generalize the above discussion we must consider signals where the cross expectation is not zero. In this regard, one of the most useful relationships is the Schwarz inequality, which can be formulated as shown in equation (A3.9):

$$\left| \int_{a}^{b} s_{1}(x) . s_{2}^{*}(x) dx \right|^{2} \leq \int_{a}^{b} |s_{1}(x)|^{2} dx \int_{a}^{b} |s_{2}(x)|^{2} dx$$
(A3.9)

where the equality only holds if $s_1(x) = k \cdot s_2(x)$ $k \in C$. Using more compact notation we can write $\int s_1(x) s_2^*(x) dx = \langle s_1 s_2^* \rangle$, and with this, equation (A3.9) can be rewritten as shown in equation (A3.10):

$$0 \le \gamma = \frac{\left|\left\langle s_1 s_2^*\right\rangle\right|}{\sqrt{\left\langle s_1 s_1^*\right\rangle \left\langle s_2 s_2^*\right\rangle}} \le 1 \tag{A3.10}$$

This ratio of integrals is called the coherence γ between signals s_1 and s_2 . From equation (A3.8) we see that the coherence is always zero for noise signals, while for polarised EM waves it follows that we can always write $E_x = kE_y$ for some complex constant k, and the coherence of polarised waves is always unity. Furthermore, as the mean phase of $\langle s_1 s_2^* \rangle$ may not be zero, it is convenient to define the complex coherence as shown in equation (A3.11):

$$\tilde{\gamma} = \gamma e^{i\phi} = \frac{\langle s_1 s_2^* \rangle}{\sqrt{\langle s_1 s_1^* \rangle \langle s_2 s_2^* \rangle}}$$
(A3.11)



Fig. A3.2 Unit circle in the complex coherence plane

This has a magnitude between 0 and 1 and a phase from 0 to 2π , hence we can represent the coherence as a point P inside the unit circle of the complex coherence plane as shown in Figure A3.2. Noise sits at the origin of this diagram, while coherent signals lie around the outer unit circle.

Coherence is a ratio of random variables and hence is a stochastic quantity, so attention must be paid to its statistics. In the most general case where correlation is allowed between the signals s_1 and s_2 , the probability density function becomes a multivariate Gaussian of the form shown in equation (A3.12):

$$\underline{u} = \begin{bmatrix} s_1 \\ s_2 \end{bmatrix} \Rightarrow p(\underline{u}) = \frac{1}{\pi^2 \det([C])} e^{-\underline{u}^{*T}[C]^{-1}\underline{u}}$$
(A3.12)

where $[C] = E(\underline{u} \cdot \underline{u}^{T*})$ is the 2 × 2 Hermitian covariance matrix defined as

$$[C] = \left\langle \begin{bmatrix} s_1 \\ s_2 \end{bmatrix} \cdot \begin{bmatrix} s_1^* & s_2^* \end{bmatrix} \right\rangle = \begin{bmatrix} \langle s_1 s_1^* \rangle & \langle s_1 s_2^* \rangle \\ \langle s_2 s_1^* \rangle & \langle s_2 s_2^* \rangle \end{bmatrix} \quad \det([C]) \ge 0 \qquad (A3.13)$$

Note that if $s_1 = ks_2$ then det([C]) = 0, and the density function must be replaced by a delta function at $\underline{u} = u_0$. The single-look density function for the phase of $s_1s_2^*$ now has the following form:

$$P(\phi) = \frac{\left(1 - \gamma^2\right) \left[\sqrt{1 - \psi^2} + \psi \left(\pi - \cos^{-1} \psi\right)\right]}{2\pi \left(1 - \psi^2\right)^{1.5}}$$
$$\psi = \gamma \cos \phi \quad -\pi \le \phi \le \pi \tag{A3.14}$$

which we note is a function of the underlying coherence γ . As the coherence reduces so the width of this distribution increases and the noise variance increases. Hence coherence and phase variance are closely related: low coherence leads to high variance, and in the limit of unit coherence the phase variance falls to zero and the phase becomes a deterministic parameter. Again, multi-look

averaging can be used to reduce the variance of the estimates for any given γ . We can therefore define the maximum likelihood estimate of [C], denoted [Z], as shown in equation (A3.15):

$$[Z] = \frac{1}{L} \sum_{j=1}^{L} \underline{u}_k \underline{u}_k^{*T}$$
(A3.15)

The matrix [Z] is then itself a random matrix with a probability distribution; the complex Wishart distribution—a function of the number of samples *L* and of the general form shown in equation (A3.16) (Lee, 1994b, 2008; Conradsen, 2003):

$$p_L([Z]) = \frac{L^{Lq} \det([Z])^{L-q} \exp(-L.Trace([C]^{-1}[Z]))}{K(L,q) \det([C])^L}$$
(A3.16)
$$K(L,q) = \pi^{0.5q(q-1)} \Gamma(L)..\Gamma(L-q+1)$$

Here q is the dimension of the complex vector \underline{u} (2 in this case, but 3 for monostatic polarimetry, 4 for bistatic, and 6 for single baseline polarimetric interferometry and so on). Equation (A3.4) is a special case of this distribution for q = 1. This distribution then leads to the following pdf for the phase as a function of the number of looks *L*:

$$P(\phi) = \frac{\Gamma\left(L + \frac{1}{2}\right)(1 - \gamma^2)^L \psi}{2\sqrt{\pi}\Gamma\left(L\right)\left(1 - \psi^2\right)^{L + \frac{1}{2}}} + \frac{(1 - \gamma^2)^L}{2\pi}F\left(L, 1 : \frac{1}{2}; \psi^2\right)$$

$$\psi = \gamma \cos\left(\phi - \phi_m\right)$$
(A3.17)

where *F* is a Gauss hypergeometric function, and ϕ_m is the mean phase. Combining these ideas, the maximum likelihood sample complex coherence is often directly used as shown in equation (A3.18):

$$\tilde{\gamma} = \frac{\sum_{i=1}^{L} s_{1i} s_{2i}^{*}}{\sqrt{\sum_{i=1}^{L} s_{1i} s_{1i}^{*}} \sqrt{\sum_{i=1}^{L} s_{2i} s_{2i}^{*}}} \begin{cases} 0 \le |\tilde{\gamma}| \le 1\\ 0 \le \arg(\tilde{\gamma}) = \hat{\phi}_{m} < 2\pi \end{cases}$$
(A3.18)

The pdf of the sample coherence magnitude $g = |\tilde{\gamma}|$ for jointly Gaussian signals can be derived analytically, and is a function of the coherence magnitude γ , the number of integrated independent samples, *L*, and the hypergeometric function *F*, as shown in equation (A3.19):

$$p(g,\gamma) = 2(L-1)(1-\gamma^2)^L g(1-g^2)^{L-2} F(L,L;1;g^2\gamma^2)$$
(A3.19)

from which the moments of order k can be deduced as shown in equation (A3.20):

$$m_{k} = \frac{\Gamma(L)\Gamma\left(1+\frac{k}{2}\right)}{\Gamma\left(L+\frac{k}{2}\right)}{}_{3}F_{2}\left(1+\frac{k}{2},L,L;L+\frac{k}{2},1;\gamma^{2}\right)\left(1-\gamma^{2}\right)^{N}$$
(A3.20)

where ${}_{p}F_{q}$ is the generalized hypergeometric function. Of particular interest is the expression for the first moment of *g*, shown in equation (A3.21):

$$E(g) = \frac{\Gamma(L)\Gamma(1+\frac{1}{2})}{\Gamma(L+\frac{1}{2})} {}_{3}F_{2}\left(\frac{3}{2}, L, L; L+\frac{1}{2}, 1; \gamma^{2}\right) \left(1-\gamma^{2}\right)^{N}$$
(A3.21)

This shows a bias towards higher coherence values, especially for low coherence with a small number of samples *L*. The variance of the estimate can also be derived from equation (A3.22) using equations (A3.20) (k = 2) and (A3.21).

$$\operatorname{var}(g) = E\left(g^2\right) - E(g)^2 \tag{A3.22}$$

Useful as these expressions are, they are difficult to interpret without detailed calculations. For this reason the Cramer–Rao (CR) lower bounds on variance of phase and coherence have also been derived (see Seymour (1994) and yield simpler expressions, as shown in equation (A3.23):

$$\operatorname{var}(\phi) > \frac{1 - \gamma^2}{2L\gamma^2} \quad \operatorname{var}(g) > \frac{(1 - \gamma^2)^2}{2L}$$
 (A3.23)

Figure A3.3 shows examples of these CR bounds as a function of coherence and increasing number of looks. Generally, the higher the coherence, the lower the number of looks required to obtain a specified variance.

Our representation of coherence inside the unit circle of Figure A3.2 is therefore rather misleading. In fact, each point P has a minimum cloud of uncertainty around it representing the Cramer–Rao bounds in radius (coherence) and polar angle (phase) fluctuations. Note that this cloud will be elliptical in shape. For a given number of looks L the phase variance is larger than the radial coherence variance. Figure A3.4 shows a schematic representation of this concept. This result underpins our distinction in Chapters 7 and 8 between coherence loci and associated coherence regions.



Fig. A3.3 Cramer–Rao bounds on fluctuations in complex coherence estimates (phase (left) and coherence amplitude (right)



Fig. A3.4 Complex coherence as a stochastic variable inside the unit circle

Having established the general properties of coherence and stochastic signals in C2, we now turn to consider the special case of depolarisation effects in polarimetry. Equation (A3.16) represents a general expression for the analysis of fluctuation statistics for coherency matrices of arbitrary dimension q. For example, q = 1 represents a scalar channel, when the distribution reduces to the gamma distribution for L > 1, with the special case of the exponential for L = 1. Dual polarised systems based on the wave coherency matrix [J]and single polarisation radar interferometry are both examples represented by q = 2. Polarimetry based on the full scattering matrix requires either q = 3for reciprocal backscatter, or q = 4 for bistatic scattering. When we consider extension to polarimetric interferometry, then the dimension increases to q =MN, where M – 1 is the number of baselines and N the number of polarisations.

In these higher-dimensional cases, analytical manipulation to find marginal distributions for phase parameters is difficult (Lopez Martinez 2005). It is then useful to employ numerical investigations based on Monte Carlo simulations using Gaussian random number generators. To see this, we start with a reference or desired MN dimensional coherency matrix Λ_{MN} . This positive semi-definite Hermitian matrix can always be expressed in terms of its eigenvalue/eigenvector decomposition, as shown in equation (A3.25):

$$[\Lambda_{MN}] = [U_{MN}] [D_{MN}] [U_{MN}]^{*T} \quad [D_{MN}] = \begin{bmatrix} \lambda_1 & 0 & 0 & 0 \\ 0 & \lambda_2 & 0 & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & 0 & \lambda_{MN} \end{bmatrix}$$
(A3.25)

This can then be used to generate a sequence of random *N*-dimensional complex sample vectors \underline{u} , all of which have a coherency matrix equal to $\mathbf{\Lambda}$ (in the limit of an infinite number of samples). We can generate such a numerical sequence using MN sets of pairs of $G(0, \sigma)$ random number generators, as shown in

equation (A3.26):

$$\underline{u} = [U_{MN}][E] = \begin{bmatrix} e_1 & 0 & 0\\ 0 & \ddots & 0\\ 0 & 0 & e_{MN} \end{bmatrix} \quad e_i = \sqrt{\lambda_i} \left\{ G_a\left(0, \frac{1}{2}\right) + iG_b(0, \frac{1}{2}) \right\}$$
$$\hat{\Lambda}_{MN} = \sum_{i=1}^{L} \underline{u} \underline{u}^{*T} \xrightarrow{L \to \infty} [\Lambda_{MN}] \quad (A3.26)$$

We start by generating two (independent) real random sequences G_a and G_b as shown, then combine them into a complex series before scaling by the square root of the appropriate eigenvalue of Λ . This process is then repeated MN times for each eigenvalue, to obtain a set of MN complex series. Finally, we introduce the complex correlations between samples by multiplying by the matrix of eigenvectors $[U_{MN}]$. The vector \underline{u} then has the property that its coherency matrix $\langle \underline{u} \cdot \underline{u}^{*T} \rangle$ converges to Γ . This provides us with a practical way to generate test sequences in polarisation statistics and depolarisation studies.

Very often in applications we make a measurement of a scattering matrix (or vector \underline{k}) and wish to determine which class it belongs to from a set of preselected reference states. This comparison process is made complicated by the stochastic nature of such measurements. For example, if \underline{k} is complex normal distributed then an individual sample may not correspond exactly to the correct class mean, and there will be some natural fluctuation. One way to deal with this is to employ a maximum likelihood (ML) approach. According to this we assign a sample to the class with the maximum probability. To do this we first need to assume a distribution (multivariate normal, for example), and then characterize each reference state by the parameters of this distribution. In the normal case this is just the covariance matrix [C], as shown in equation (A3.27):

$$\underline{u} = \begin{bmatrix} s_1 \\ \vdots \\ s_q \end{bmatrix} \Rightarrow p(\underline{u}) = \frac{1}{\pi^q \det([C])} e^{-\underline{u}^{s^T}[C]^{-1}\underline{u}}$$
(A3.27)

Each class is then characterized by a $q \times q$ class covariance matrix [C_i], which we must calculate or measure before the comparison takes place. We then take the measured vector \underline{k} and compare it to all the class matrices. Geometrically this reduces to a distance measure between the sample vector and class covariance. As the normal distribution involves the exponential function, it is common to consider distances based on the so-called log-likelihood function, obtained from the normal distribution by taking the natural logarithm, as shown in equation (A3.28), where we have used the cyclic property of the trace operation to simplify the centre term.

$$-\ln|C_i| - Tr(C_i^{-1}\underline{kk}^{*T}) - q\ln\pi$$
 (A3.28)

From this we can define a non-negative distance measure such that we assign \underline{k} to the class with the shortest distance d, defined from equation (A3.28) by

ignoring elements that do not depend on the class, as shown in equation (A3.29):

$$d\left(\underline{k}, C_{i}\right) = \ln|C_{i}| + Tr(C_{i}^{-1}\underline{k}\underline{k}^{*T})$$
(A3.29)

This is formally a measure of the 'closeness' of \underline{k} to class *i*. It forms the basis for image classification and hypothesis testing in radar polarimetry and interferometry (Lee, 2008).

In the depolarising case we may wish to compare not a single \underline{k} vector but an average coherency or covariance matrix C itself. A distance measure for this case can be obtained in a similar fashion to equation (A3.29), but starting from the complex Wishart distribution of equation (A3.16), and again forming the log-likelihood function and ignoring constant terms to obtain equation (A3.30):

$$d(C, C_i) = \ln |C_i| + tr(C_i^{-1}C)$$
(A3.30)

Note that since the Mueller matrix [M] can be mapped 1–1 with the scattering coherency matrix [T], which itself is unitarily similar to [C], the metric in equation (A3.30) is invariant to use of [C] or [T]. In this way we can also provide a statistical distance measure between experimental Mueller matrices.

Bibliography

- Abhyankar, K. D. and Fymat, A. L. (1969). Relations between the elements of the phase matrix for scattering. *Journal of Math. Phys.*, 10, 1935–1938.
- Ablitt, B. P., Hopcraft, K. I., Turpin, K. D., Chang, P. C. Y. and Walker, J. G. (1999). Imaging and multiple scattering through media containing optically active particles. *Waves in Random Media*, 9, 561–572.
- Ablitt, B. P. (2000). *Characterisation of Particles and their Scattering Effects on Polarized Light*. PhD thesis, University of Nottingham, UK.
- Ainsworth, T. L., Ferro-Famil, L. and Lee, J. S. (2006). Orientation angle preserving *a posteriori* polarimetric SAR calibration. *IEEE Transactions* on Geoscience and Remote Sensing, 44 (4), 994–1003.
- Ainsworth, T. L., Preiss, M., Stacy, N., Nord, M. and Lee, J. S. (2007). Analysis of compact polarimetric SAR imaging modes. Proceedings of the Third ESA Workshop on Polarimetry and Polarimetric Interferometry, POLInSAR 2007, Frascati, Italy, January 2007. http://earth.eas/int/workshops/polinsar2007/
- Allain, S. (2003). Caractérisation d'un Sol nu à partir de données SAR Polarimétriques: Etude Multi-fréquentielle et Multi-résolutions. PhD thesis, University of Rennes, France.
- Anderson, D. G. M. and Barakat, R. (1994). Necessary and sufficient conditions for a Mueller matrix to be derivable from a Jones matrix. *J. Opt. Soc. Am. A*, **11** (8), 2305–2319.
- Askne, J., Dammert, P. B., Ulander, L. M. and Smith, G. (1997). C-band repeat pass interferometric SAR observations of the forest. *IEEE Transactions* on Geoscience and Remote Sensing, 35, 25–35.
- Askne, J., Santoro, M., Smith, G. and Fransson, J. E. S. (2003). Multitemporal repeat-pass SAR interferometry of boreal forests. *IEEE Transactions on Geoscience and Remote Sensing*, **41**, 1540–1550.
- Askne, J. and Santoro, M. (2007). Selection of forest stands for stem volume retrieval from stable ERS tandem InSAR observations. *IEEE Geoscience and Remote Sensing Letters*, **4**, 46–50.
- Attema, E. P. and Ulaby, F. T. (1978). Vegetation modeled as a water cloud. *Radio Science*, **13**, 357–364.
- Azzam, R. M. (1978). Propagation of partially polarised light. *J. Opt. Soc. Am.*, **68**, 1756–1767.
- Azzam, R. M. A. and Bashara, N. M. (1987). *Ellipsometry and Polarized Light*. North–Holland.

- Ballester-Berman, J. D., Lopez-Sanchez, J. M. and Fortuny-Guasch, J. (2005). Retrieval of biophysical parameters of agricultural crops using polarimetric SAR interferometry. *IEEE Transactions on Geoscience and Remote Sensing*, **43** (4), 683–694.
- Ballester-Berman, J. D. and Lopez-Sanchez, J. M. (2007). Coherence loci for a homogeneous volume over a double-bounce ground return. *IEEE Geoscience and Remote Sensing Letters*, **4** (2), 317–321.
- Bamler, R. (1992). A comparison of range-Doppler and wavenumber domain SAR focusing algorithm. *IEEE Transactions on Geoscience and Remote Sensing*, **30**, 706–713.
- Bamler, R. and Hartl, P. (1998). Synthetic aperture radar interferometry. *Inverse Problems*, **14**, R1–R54.
- Barakat, R. (1981). Bilinear constraints between the elements of the 4×4 Mueller–Jones matrix of polarization theory. *Opt. Comms.*, **38**, 159–161.
- Barakat, R. (1987). Conditions for the physical realisability of polarization matrices characterising passive systems. J. Mod. Optics, 34, 1535– 1544.
- Bates, A. P., Hopcraft, K. I. and Jakeman, E. (1998). Non-Gaussian fluctuations of Stokes parameters in scattering by small particles. *Waves in Random Media*, 8, 235–253.
- Baum, C. and Kritikos H. N. (eds.) (1995). *Electromagnetic Symmetry*. Taylor and Francis, Washington.
- Bessette, L. A. and Ayasli, S. (2001). Ultra wide band P-3 and Carabas II foliage attenuation and backscatter analysis. *Proceedings of IEEE Radar Conference*, 357–362.
- Bickel, S. H. and Bates, R. H. T. (1965). Effects of magneto-ionic propagation on the scattering matrix. *Proc. IEEE*, **53** (8), 1089–1091.
- Bickel, W. S. and Bailey W. M. (1985). Stokes vectors, Mueller matrices and polarized scattered light. *American Journal of Physics*, **53**, 468–478.
- Bicout, D. and Brosseau C. (1992). Multiply scattered waves through a spatially random medium: entropy production and depolarization. *J. Phys. I. France*, 2, 2047–2063.
- Boerner, W. M. (1981). Polarization dependence in electromagnetic inverse problems. *IEEE Trans. Antennas and Propagation*, AP-29, 262–274.
- Boerner, W. M. (ed.) (1992). *Direct and Inverse Methods in Radar Polarimetry, Parts 1 and 2*. NATO ASI Series C: Mathematical and Physical Sciences, Vol. 350, Kluwer.
- Borgeaud, M. and Noll, J. (1994). Analysis of theoretical surface scattering models for polarimetric microwave remote sensing of bare soils. *International Journal of Remote Sensing*, **15** (14), 2931–2942.
- Born, M. and Wolf E. (1998). *Principles of Optics*, Chapters 1 and 10. Pergamon Press, sixth edition.
- Brosseau, C. (1990). Analysis of experimental data for Mueller polarization matrices. OPTIK, 85, 83–86.
- Brosseau, C. and Bicout, D. (1994). Entropy production in multiple scattering of light by a spatially random medium. *Phys. Rev. E*, **50**, 4997–5005.
- Brosseau, C. (1998). Fundamentals of Polarized Light: a Statistical Approach. Wiley.

- Byrne, J. (1971). Classification of electron and optical polarization transfer matrices. J. Phys. B, 4, 940–953.
- Cafforio, C., Pratti, C. and Rocca, F. (1991). SAR data focusing using seismic migration techniques. *IEEE Trans. Aerospace and Electronic Systems*, 27, 194–205.
- Cameron, W. L., Youssef, N. N. and Leung, L. K. (1996). Simulated polarimetric signatures of primitive geometrical shapes. *IEEE Transactions on Geoscience and Remote Sensing*, 34 (3), 793–803.
- Cartan, E. (1966). The Theory of Spinors. Dover Press.
- Chandrasekhar, S. (1960). Radiative Transfer. Dover Press.
- Chen, H. C. (1985). *Theory of Electromagnetic Waves: a Coordinate-Free Approach*. McGraw-Hill.
- Cloude, S. R. (1985). Radar target decomposition theorems. *IEEE Letters*, **21** (1), 22–24.
- Cloude, S. R. (1986). Group theory and polarization algebra. *OPTIK*, **75** (1), 26–36.
- Cloude, S. R. (1989). Physical realisability of matrix operators in polarimetry. SPIE, 1166, Polarization Considerations for Optical Systems II, pp. 177–185.
- Cloude, S. R. (1995a). An Introduction to Electromagnetic Wave Propagation and Antennas. UCL Press.
- Cloude, S. R. (1995b). Lie groups in EM wave propagation and scattering. In Baum, C. and Kritikos, H. N. (eds.), *Electromagnetic Symmetry*, Chapter 2. Taylor and Francis, Washington.
- Cloude, S. R. and Pottier, E. (1995c). The concept of polarisation entropy in optical scattering. *Optical Engineering*, **34** (6), 1599–1610.
- Cloude, S. R. and Pottier, E. (1996). A review of target decomposition theorems in radar polarimetry. *IEEE Transactions on Geoscience and Remote Sensing*, **34** (2), 498–518.
- Cloude, S. R. and Pottier, E. (1997a). An entropy based classification scheme for land applications of polarimetric SAR. *IEEE Transactions on Geoscience and Remote Sensing*, **35** (1), 68–78.
- Cloude, S. R. and Papathanassiou, K. P. (1997b). Polarimetric optimisation in radar interferometry. *Electronics Letters*, **33** (13), 1176–1178.
- Cloude, S. R. and Papathanassiou, K. P. (1998). Polarimetric SAR interferometry. *IEEE Transactions on Geoscience and Remote Sensing*, GRS-36 (5), 1551–1565.
- Cloude, S. R., Fortuny, J., Lopez, J. M. and Sieber, A. J. (1999). Wide band polarimetric radar inversion studies for vegetation layers. *IEEE Transactions on Geoscience and Remote Sensing*, 37/2 (5), 2430–2442.
- Cloude, S. R., Papathanassiou, K. P. and Boerner, W. M. (2000a). The remote sensing of oriented volume scattering using polarimetric radar interferometry. *Proceedings of International Symposium on Antennas and Propagation, ISAP 2000, Fukuoka, Japan*, 549–552.
- Cloude, S. R., Papathanssiou, K. P. and Reigber, A. (2000b). Polarimetric SAR interferometry at P band for vegetation structure extraction. *Proceedings of the Third European SAR Conference, EUSAR 2000, Munich, Germany*, 249–252.

- Cloude, S. R., Papathanssiou, K. P. and Boerner, W. M. (2000c). A fast method for vegetation correction in topographic mapping using polarimetric radar interferometry. *Proceedings of the Third European SAR Conference*, *EUSAR 2000, Munich, Germany*, 261–264.
- Cloude, S. R. (2001a). A new method for characterising depolarisation effects in radar and optical remote sensing. *Proceedings of the IEEE International Geoscience and Remote Sensing Symposium, IGARSS 2001, Sydney, Australia*, **2**, 910–912.
- Cloude, S. R., Papathanassiou, K. P. and Pottier, E. (2001b). Radar polarimetry and polarimetric interferometry. *IEICE Transactions on Electronics*, E84-C (12), 1814–1822.
- Cloude, S. R., Woodhouse, I. H., Hope, J., Suarez Minguez, J. C., Osborne, P. and Wright G. (2001c). The Glen Affric Radar Project: forest mapping using dual baseline polarimetric radar interferometry. ESA Symposium on Retrieval of Bio and Geophysical Parameters from SAR for Land Applications, University of Sheffield, England, 333–338.
- Cloude, S. R. and Corr, D. G. (2002a). A new parameter for soil moisture estimation. *Proceedings of the IEEE International Geoscience and Remote Sensing Symposium, IGARSS 2002, Toronto, Canada*, **1**, 641–643.
- Cloude, S. R. (2002b). Helicity in radar remote sensing. *Proceedings of the IEEE International Geoscience and Remote Sensing Symposium, IGARSS* 2002, *Toronto, Canada*, **1**, 411–413.
- Cloude, S. R. and Papathanassiou, K. P. (2003). A 3-stage inversion process for polarimetric SAR interferometry. *IEEE Proceedings, Radar, Sonar and Navigation*, **150** (03), 125–134.
- Cloude, S. R., Corr, D. G. and Williams, M. L. (2004). Target detection beneath foliage using polarimetric SAR interferometry. *Waves in Random Media*, 14 (2), S393–S414.
- Cloude, S. R. and Williams, M. L. (2005a). The negative alpha filter: a new processing technique for polarimetric SAR interferometry. *IEEE Geoscience and Remote Sensing Letters*, **2**, 187–191.
- Cloude, S. R. (2005b). On the status of bistatic polarimetry theory. *Proceedings* of *IEEE Geoscience and Remote Sensing Symposium, IGARSS 2005, Seoul, South Korea,* **3**, 2003–2006.
- Cloude, S. R. (2006a). Information extraction in bistatic polarimetry. *Proceedings of the Sixth European SAR Conference, EUSAR 06, Dresden, Germany.*
- Cloude, S. R. (2006b). Polarization coherence tomography, *Radio Science*, **41**, RS4017.
- Cloude, S. R. (2007a). Dual baseline coherence tomography. *IEEE Geoscience* and Remote Sensing Letters, **4** (1), 127–131.
- Cloude, S. R. (2007b). The dual polarization H/alpha decomposition. *Proceedings of the Third ESA Workshop on Polarimetry and Polarimetric Interferometry, POLInSAR 2007, Frascati, Italy.*
- Colin, E., Titin-Schnaider, C. and Tabbara, W. (2006). An interferometric coherence optimization method in radar polarimetry for high-resolution imagery. *IEEE Transactions on Geoscience and Remote Sensing*, 44 (1), 167–175.

- Colin, E. (2005). Apport de la Polarimétrie à l'Interférométrie Radar pour l'Estimation des Hauteur's de Cibles et de Paramétres de Forêt, PhD thesis, Université de Paris.
- Collet, E. (1993). Polarized Light. Marcel Dekker, New York.
- Collin, R. E. (1985). Antennas and Radiowave Propagation. McGraw-Hill.
- Conradsen, K., Nielsen, A. A., Schou, J. and Skriver, H. (2003). A test statistic in the complex Wishart distribution and its application to change detection in polarimetric SAR data. *IEEE Transactions on Geoscience and Remote Sensing*, **41** (1). 4–19.
- Cornwell, J. F. (1984). *Group Theory in Physics*, Vol. 1, 'Techniques in physics', 7. Academic Press.
- Curlander, J. C. and McDonough, R. N. (1991). *Synthetic Aperture Radar: Systems and Signal Processing*. Wiley Series in Remote Sensing.
- Dall, J., Papathanassiou, K. P. and Skriver, H. (2003). Polarimetric SAR interferometry applied to land ice: first results. *Proceedings of the IEEE Geoscience and Remote Sensing Symposium, IGARSS '03, Toulouse, France*, 3, 1432–1434.
- Deschamps, G.A. (1951). Geometrical representation of plane polarized waves. *Proc. IRE*, **39**, 540.
- Dobson, M. C., Ulaby, F. T., Hallikainen, M. and El-Rayes, M. A. (1085). Microwave dielectric behaviour of wet soil: II Four component dielectric mixing models. *IEEE Transactions on Geoscience and Remote Sensing*, 23, 35–46.
- Dong, Y., Forster, B. C. and Ticehurst, C. (1998). A new decomposition of radar polarization signatures. *IEEE Transactions on Geoscience and Remote Sensing*, GRS-36, 933–939.
- Dubois, P. C., van Zyl, J. J. and Engman, T. (1005). Measuring soil moisture with imaging radars. *IEEE Transactions on Geoscience and Remote Sensing*, GE-33, 916–926.
- Ferro-Famil, L. and Pottier, E. (2000). Description of dual frequency polarimetric data using Gell–Mann parameter set. *Electronics Letters*, **36** (19), 1646–1647.
- Ferro-Famil, L., Pottier, E. and Lee, J. S. (2001). Unsupervised classification of multifrequency and fully polarimetric SAR images based on the H/A/Alpha–Wishart classifier. *IEEE Transactions on Geoscience and Remote Sensing*, **39** (11), 2332–2342.
- Ferro-Famil, L., Reigber, A., Pottier, E. and Boerner, W. M. (2003). Scene characterization using subaperture polarimetric SAR data. *IEEE Transactions* on Geoscience and Remote Sensing, **41** (10), Part 1, 2264–2276.
- Ferro-Famil, L. and Neumann, M. (2008). Recent advances in the derivation of POLInSAR statistics: study and applications. *Proceedings of the Seventh European Conference on Synthetic Aperture Radar (EUSAR)*, *Friedrichshafen, Germany*, 2, 143–146.
- Flynn T., Tabb, M. and Carande, R. (2002). Coherence region shape estimation for vegetation parameter estimation in POLINSAR. *Proceedings of IGARSS 2002, Toronto, Canada*, V 2596–2598.
- Franceschetti, G. and Linari, R. (1999). *Synthetic Aperture Radar Processing*. Chapter 4. CRC Press.

- Freeman, A. (1992). SAR calibration: a review. *IEEE Transactions on Geoscience and Remote Sensing*, GE-30(6), 1107–1121.
- Freeman, A. and Durden, S. L. (1998). A three component model for polarimetric SAR data. *IEEE Transactions on Geoscience and Remote Sensing*, GE-36, 963–973.
- Freeman, A. (2004). Calibration of linearly polarized polarimetric SAR data subject to Faraday rotation. *IEEE Trans.*, GRS-42 (8), 1617–1624.
- Freeman, A. (2007). Fitting a two component scattering model to polarimetric SAR data. *IEEE Trans*. GRS-42 (8), 2583–2592.
- Fry, E. S. and Kattawar, G. W. (1981). Relationships between elements of the Stokes matrix. *Applied Optics*, 20, 2811–2814.
- Fung, A. K., Li, Z. and Chen, K. S. (199). Backscattering from a randomly rough dielectric surface. *IEEE Transactions on Geoscience and Remote Sensing*, **30**, 356–369.
- Gatelli, F., Monti Guarnieri, A., Parizzi, F., Pasquali, P., Prati, C. and Rocca, F. (1994). The wavenumber shift in SAR interferometry. *IEEE Trans.*, GRS-32, 855–865.
- Gazdag, J. and Sguazzero, P. (1984). Migration of seismic data. *Proceedings* of the IEEE, **72**, 1302–1315.
- Georgi, H. (1999). Lie Algebras in Particle Physics. Perseus Books.
- Gershenfeld, N. (1999). *The Nature of Mathematical Modeling*. Cambridge University Press.
- Gil, J. J. and Bernabeu, E. (1985). A depolarization criterion in Mueller matrices. *Optica Acta*, **32**, 259–261.
- Gil, J. J. and Bernabeu, E. (1986). Depolarization and polarization indices of an optical system. *Optica Acta*, **33**, 185–189.
- Girgel, S. S., (1991). Structure of the Mueller matrices of depolarised optical systems. *Sov. Phys. Crystallogr.*, **36**, 890–891.
- Giuli. D. (1986). Polarization diversity in radars. *Proceedings of the IEEE*, **74**, 245–269.
- Givens, C. R. and Kostinski, A. B. (1993). A simple necessary and sufficient condition on physical realizable Mueller matrices. *J. Mod. Opt.*, **40**, 471–481.
- Goldstein, H. (1980). *Classical Mechanics*, second edition. Addison-Wesley.
- Graham, R. (1974). Synthetic interferometric radar for topographic mapping. *Proceedings of the IEEE*, **62**, 763–768.
- Graves, C. D. (1956). Radar polarization power scattering matrix. *Proceedings* of the IRE, 44 (2), 248–252.
- Hagberg, J. O., Ulander, L. and Askne, J. (1995). Repeat-pass SAR interferometry over forested terrain. *IEEE Transactions on Geoscience and Remote Sensing*, **33** (2), 331–340.
- Hajnsek, I., Papathanassiou, K. P. and Cloude, S. R. (2001). Removal of additive noise in polarimetric eigenvalue processing. *Proceedings of the IEEE Symposium on Geoscience and Remote Sensing, IGARSS '01*, 6, 2778– 2780.

- Hajnsek, I., Pottier, E. and Cloude, S. R. (2003). Inversion of surface parameters from polarimetric SAR. *IEEE Transactions on Geoscience and Remote Sensing*, **41**, 727–744.
- Hajnsek, I., Kugler, F., Lee, S. K. and Papathanassiou, K. P. (2008). Tropical forest parameter estimation by means of POLInSAR: the INDREX-II Campaign. *IEEE Transactions on Geoscience and Remote Sensing*, 47, 481–493.
- He, C. and Watson, G. A. (1997). An algorithm for computing the numerical radius. *IMA J. Numer. Anal.*, **17**, 329–342.
- Hecht, E. and Zajac, A. (1997). Optics. Third edition. Addison-Wesley.
- Hopcraft, K. I. and Smith, P. R. (1992). An introduction to electromagnetic inverse scattering. *Developments in EM Theory and Applications*, **7**.
- Hovenier, J. W. (1994). Structure of a general pure Mueller matrix. *Applied Optics*, **33**, 8318–8324.
- Hovenier, J. W. and van der Mee, C. V. M. (1996). Testing scattering matrices: a compendium of recipes. *Journal of Quantitative Spectroscopy and Radiative Transfer*, **55**, 649–661.
- Hovenier, J. W., van der Mee, C.V.M. and Domke, H. (2004). *Transfer of Polarized Light in Planetary Atmospheres: Basic Concepts and Practical Methods*. Kluwer Academic Publishers, Astrophysics and Space Science Library, Vol. 318.
- Hunt, B. J. (1991). The Maxwellians. Cornell University Press.
- Huynen, J. R. (1970). *Phenomenological Theory of Radar Targets*, PhD thesis, Technical University, Delft, Netherlands.
- Huynen, J. R. (1987). Phenomenological theory of radar targets. In Uslenghi, P. L. E. (ed.), *Electromagnetic Scattering*, Academic Press, New York.
- Imhoff, M. L. (1995). Radar backscatter and biomass saturation: ramifications for global biomass inventory. *IEEE Transactions on Geoscience* and Remote Sensing, 33, 511–518.
- Iniesta, J. C. del Toro (2003). Introduction to Spectropolarimetry. Cambridge.
- Ioannidis, G. A. and Hammers, D. E. (1979). Optimum antenna polarizations for target discrimination in clutter. *IEEE Trans. Antennas and Propagation*, AP-27, 357–363.
- Ishimaru, A. (1991). *Electromagnetic Wave Propagation, Radiation and Scattering*. Prentice Hall International.
- Jackson, J. D. (1999). Classical Electrodynamics. Third edition. Wiley.
- Jin, Y. Q. and Cloude, S. R. (1994a). Numerical eigenanalysis of the coherency matrix for a layer of random non-spherical scatterers. *IEEE Transactions on Geoscience and Remote Sensing*, **32**., 1179–1185.
- Jin, Y. Q. (1994b). *Electromagnetic Scatering Modelling for Quantitative Remote Sensing*. World Scientific Publishing.
- Jones, D. S. (1989). *Acoustic and Electromagnetic Waves*. Oxford Science Publications.
- Jones, R. C. (1941). New calculus for the treatment of optical systems. J. Opt. Soc. Am., **31**, 488–493.
- Jones, R. C. (1948). New calculus for the treatment of optical systems, VII: Properties of the N-matrices. J. Opt. Soc. Am., 38, 671–685.

- Kampes, B. M. (2006). *Radar Interferometry: Persistent Scatterer Technique*. Kluwer.
- Kennaugh, E. M. (1952). Polarization Properties of Radar Reflections. MSc thesis, Electro-Science Laboratory, Ohio State University.
- Kim, K., Mandel, L. and Wolf, E. (1987). Relationship between Jones and Mueller matrices for random media. *Journal of Opt. Soc. Am.* A, 4, 433–437.
- Kimura, H., Mizuno, T., Papathanassiou, K. P. and Hajnsek, I. (2004). Improvement of polarimetric SAR calibration based on the Quegan algorithm. *Proceedings of the IEEE IGARSS 04 Symposium*, 1.
- Kong, J. A. (1985). Electromagnetic Wave Theory. Wiley.
- Kong, J. A. (ed.) (1990). Polarimetric remote sensing. Progress in Electromagnetics Research, PIER 3. Elsevier.
- Konnen, G. P. (1985). Polarized Light in Nature. Cambridge University Press.
- Kostinski, A. B. and Boerner, W. M. (1986). On foundations of radar polarimetry. *IEEE Trans. Antennas and Propagation*, AP-34, 1395–1404.
- Krieger, G., Papathanassiou, K. P. and Cloude, S. R. (2005). Spaceborne polarimetric SAR interferometry: performance analysis and mission concepts. *EURASIP Journal of Applied Signal Processing*, 20, 3272–3292.
- Krieger, G., Moreira, A., Fiedler, H., Hajnsek, I., Werner, M., Younis, M. and Zink, M. (2007). TanDEM-X: a satellite formation for high-resolution SAR interferometry. *IEEE Transactions on Geoscience and Remote Sensing*, 45, 3317–3341.
- Krogager, E. (1993). Aspects of Polarimetric Radar Imaging. PhD thesis, Technical University of Denmark.
- Krogager, E. (1992). Decomposition of the Sinclair matrix into fundamental components with applications to high resolution radar imaging. In Boermer, W. M. *et al.* (eds.), *Direct and Inverse Methods in Radar Polarimetry*, 2, 1459–1478. Kluwer Academic Publishers.
- Lakhtakia, A., Varadan, V. V. and Varadan, V. K. (1989). *Time Harmonic EM Fields in Chiral Media*. Springer.
- Le, C. T. C., Ishimaru, A., Kuga, Y. and Hae Yea, J. (1998). Angular memory and frequency interferometry for mean height profiling of a rough surface. *IEEE Transactions on Geoscience and Remote Sensing*, **36**, 61–67.
- Lee, J. S. (1994a). Speckle filtering of SAR images: a review. *Remote Sensing Reviews*, **8**, 313–340.
- Lee, J. S., Hoppel, K. W., Mango, S. A. and Miller, A. (1994b). Intensity and phase statistics of multi-look polarimetric and interferometric SAR imagery. *IEEE Transactions on Geoscience and Remote Sensing*, GE-32, 1017–1028.
- Lee, J. S., Grunes, M. R., Ainsworth, T. L., Du, L. J., Schuler, D. L. and Cloude, S. R. (1999). Unsupervised classification using polarimetric decomposition and the complex Wishart distribution. *IEEE Transactions* on Geoscience and Remote Sensing, **37**/1 (5), 2249–2259.
- Lee, J. S., Schuler, D. L. and Ainsworth, T. L. (2000). Polarimetric SAR sata compensation for terrain azimuth slope variation. *IEEE Transactions on Geoscience and Remote Sensing*, 38/5, 2153–2163.
- Lee, J. S., Schuler, D. L., Ainsworth, T. L., Krogager, E., Kasilingam, D. and Boerner, W.M. (2002). On the estimation of radar polarization orientation

shifts induced by terrain slopes. *IEEE Transactions on Geoscience and Remote Sensing*, **40**, 30–41.

- Lee, J. S. and Pottier, E. (2008). Polarimetric radar imaging: from basics to applications. *Optical Science and Engineering Series*, **143**, CRC Press.
- Li, R. C. (1994). *Relations between the Field of Values of a Matrix and those of its Schur Complements*. Report No. UCB//CSD-94-849, Computer Science Division, University of California at Berkeley.
- Lopez-Martinez, C., Pottier, E. and Cloude, S. R. (2005). Statistical assessment of eigenvector-based target decomposition theorems in radar polarimetry. *IEEE Transactions on Geoscience and Remote Sensing*, **43** (9), 2058– 2074.
- Lopez, J. M., Fortuny, J., Cloude, S. R. and Sieber, A. J. (2000). Indoor polarimetric radar measurements on vegetation samples at L, C, S and X bands. *Journal of Electromagnetic Waves and Applications*, 14 (2), 205–231.
- Lopez-Sanchez, J. M., Ballester-Berman, J. D. and Fortuny-Guasch, J. (2006). Indoor wide-band polarimetric measurements on maize plants: a study of the differential extinction coefficient. *IEEE Transactions on Geoscience and Remote Sensing*, **44** (4), 758–767.
- Lopez-Sanchez, J. M., Ballester-Berman, J. D. and Marquez-Moreno, Y. (2007). Model limitations and parameter-estimation methods for agricultural applications of polarimetric SAR interferometry. *IEEE Transactions on Geoscience and Remote Sensing*, **45** (11), Part 1, 3481–3493.
- Lu, S. Y. and Chipman, R. A. (1994). Homogeneous and inhomogeneous Jones matrices. JOSA A, 11 (2), 766–773.
- Lu, S. Y. and Chipman, R. A. (1996). Interpretation of Mueller matrices based on polar decomposition. *JOSA* A, **13** (5), 1106–1113.
- Ludwig, A. (1973). Definition of cross polarisation. *IEEE Transactions on Antennas and Propagation*, AP-21, 116–119.
- Luneburg, E. (1996). Polarimetry: a revision of basic concepts. In Cloude, S. R., Serbest, A. H. (eds.), *Direct and Inverse Electromagnetic Scattering*. Pitman Research Notes in Mathematics, Vol. 361, 257–275. Longman Scientific and Technical.
- Luneburg, E. and Cloude, S. R. (1997). Optimisation procedures for bistatic scattering. *SPIE Proceedings on Wideband Interferometric Sensing and Imaging Polarimetry*, **3120**.
- Macintosh, F. C., Zhu, J. X., Pine, D. J. and Weitz, D. A. (1989). Polarization memory of multiply scattered light. *Phys. Rev. B*, **40** (13), 9342–9345.
- Mattia, F., Le Toan, T., Souyris, J. C., De Carolis, C., Floury, N., Posa, F. and Pasquariello, N. G. (1997). The effect of surface roughness on multifrequency polarimetric SAR data. *IEEE Transactions on Geoscience and Remote Sensing*, **35** (4), 954–966.
- Mendez, E. R. and O'Donnell, K. A. (1987). Observation of depolarization and backscattering enhancement in light scattering for Gaussian random surfaces. *Opt. Comm.*, **61**, 91–95.
- Mengi, E. and Overton, M. L. (2005). Algorithms for the computation of the pseudospectral radius and the numerical radius of a matrix. *IMA J. Numerical Analysis*, **25**, 648–669.
- Mensa, D. L. (1991). *High Resolution Radar Cross-Section Imaging*. Artech House.

- Mette, T., Papathanassiou, K. P. and Hajnsek, I. (2004). Biomass estimation from POLInSAR over heterogeneous terrain. *Proceedings of IEEE Geoscience and Remote Sensing Symposium, IGARSS 2004, Anchorage, Alaska, 20–24 September 2004.*
- Mette, T. (2007). Forest Biomass Estimation from Polarimetric SAR Interferometry. DLR Research Report 2007-10.
- Mishchenko, M. I. (1992). Enhanced backscattering of polarized light from discrete random media: calculations in exactly the backscattering direction. J. Opt. Soc. Am. (A), 9, 978–982.
- Mishchenko, M. I. and Hovenier, J. W. (1995). Depolarization of light backscattered by randomly oriented nonspherical particles. *Optics Letters*, **20** (12), 1356–1359.
- Mishchenko, M., Hovenier, J. W. and Travis, L.D. (2000). *Light Scattering by Nonspherical Particles: Theory, Measurements and Applications.* Academic Press.
- Mishchenko, M. I., Travis, L. D. and Lacis, A. A. (2006). *Multiple Scattering* of Light by Particles: Radiative Transfer and Coherent Backscattering. Cambridge.
- Mishchenko, M. I., Liu, L., Mackowski, D. W., Cairns, B. and Videen, G. (2007). Multiple scattering by random particulate media: exact 3D results. *Optics Express*, **15** (6, 19), 2822–2836.
- Misner, C. W., Thorne, K. S. and Wheeler, J. A. (1992). *Gravitation*. W. H. Freeman and Co.
- Mott, H. (1992). Antennas for Radar and Communications. Wiley.
- Mott, H. (2007). Remote Sensing with Polarimetric Radar. Wiley Interscience.
- Murnaghan, F. D. (1932). On the field of values of a square matrix. *Proc. N. A. S. Mathematics*, 246–248.
- Murnaghan, F. D. (1962). The Unitary and Rotation Groups. Spartan Books.
- Nelander, A. (1995). Analysis of wide band polarimetric radar. *Proceedings* of the third International Workshop on Radar Polarimetry (JIPR '95), IRESTE, University of Nantes, France, 89–98.
- Neumann, M., Ferro-Famil, L. and Reigber, A. (2008). Multibaseline polarimetric SAR interferometry coherence optimization. *IEEE Geoscience and Remote Sensing Letters*, **5**, (1), 93–97.
- Nghiem, S. V., Yueh, S. H., Kwok, R. and Li, F. K. (1992). Symmetry properties in polarimetric remote sensing. *Radio Science*, **27** (5), 693–711.
- Novak, L. M., Sechtin, M. B. and Cardullo, M. J. (1989). Studies of target detection algorithms which use polarimetric radar data. *IEEE Trans. Aerospace* and Electronic Systems, AES-25, 15–165.
- Nye, J. F. (1999). Natural Focusing and Fine Structure of Light. IoP Publishing.
- Oh, Y., Sarabandi, K. and Ulaby, F. T. (1992). An empirical model and an inversion technique for radar scattering from bare soil surfaces. *IEEE Transactions on Geoscience and Remote Sensing*, GE-30(2), 370–381.
- O'Neill, E. L. (1991). Introduction to Statistical Optics. Dover Press.
- Pancharatnam, S. (1956). Generalised theory of interference and its applications. I: Coherent pencils. *Proc. Indian Acad. Sci.* A, 44, 247–262.

- Papathanassiou, K. P. and Cloude, S. R. (1997). Polarimetric effects in repeatpass interferometry. *Proceedings IGARSS 97, Singapore, 3–8 August* 1997, 1926–1928.
- Papathanassiou, K. P. and Zink, M. (1998a). Polarimetric calibration of the airborne experimental SAR system of DLR. Proceedings of European SAR Conference, EUSAR 1998, Friedrichshafen, Germany.
- Papathanassiou, K. P., Reigber, A., Scheiber, R., Horn, R., Moreira, A. and Cloude S. R. (1998b). Airborne polarimetric SAR interferometry. *Proceedings of IEEE Symposium on Geoscience and Remote Sensing* (IGARSS), Seattle, USA, July 6–10.
- Papathanassiou, K. P. and Cloude, S. R. (2001). Single baseline polarimetric SAR interferometry. *IEEE Transactions on Geoscience and Remote Sensing*, GRS-39/11, 2352–2363.
- Papathanassiou, K. P. and Cloude, S. R. (2003). The effect of temporal decorrelation on the inversion of forest parameters from POLInSAR data. *Proceedings of IEEE International Geoscience and Remote Sensing Symposium (IGARSS 2003), Toulouse, France, July 21–25.*
- Papathanassiou, K. P., Cloude, S. R., Liseno, A., Mette, T. and Pretzsch, H. (2005). Forest height estimation by means of polarimetric SAR interferometry: actual status and perspectives. Proceedings of the Second ESA POLInSAR Workshop, Frascati, Italy, January 2005. http://earth.esa.int/workshops/polinsar2005/
- Pascual, C., Gimeno-Nieves, E. and Lopez-Sanchez, J. M. (2002). The equivalence between the polarisation subspace method (PSM) and coherence optimisation in polarimetric radar interferometry. *Proceedings of the Fourth European Synthetic Aperture Radar Conference, EUSAR 2002*, 589–592.
- Penrose, R. and Rindler, W. (1984). Spinors and Space-Time, Volume 1: Two Spinor Calculus and Relativisitic Fields. Cambridge University Press.
- Perrin, F. (1942). Polarization of light scattered by isotropic opalescent media. *Journal of Chemical Physics*, **10**, 415–427.
- Poelman, A. J. and Hilgers, C. J. (1991). Effectiveness of multinotch logicproduct polarisation filters in radar for countering rain clutter. *IEE Proceedings F, Radar and Signal Processing*, **138**, 427–437.
- Poincaré, H. (1997). *Theorie Mathematique de la Lumiere II*. Chapter 12. Paris, 1892.
- Pottier, E. and Cloude, S. R. (1997). Application of the H-A-α polarimetric decomposition theorem for land classification. *SPIE International Symposium on Optical Science Engineering and Instrumentation, Wideband Interferometric Sensing and Imaging Polarimetry, San Diego, California, USA, 27 July–1 August 1997.*
- Praks, J., Kugler, F., Papathanassiou, K. P., Hajnsek, I. and Hallikainen, M. (2007) Height estimation of boreal forest: interferometric model based inversion at L and X bands versus HUTSCAT profiling scatterometer. *IEEE Geoscience and Remote Sensing Letters*, 4, 466–470.
- Press, W. H., Teukolsky, S. A., Vetterling, W. T. and Flannery, B. P. (2007). *Numerical Recipes 3rd Edition: The Art of Scientific Computing*. Cambridge University Press.

- Priest, R. G. and Germer, T. A. (2000). Polarimetric BRDF in the microfacet model, theory and measurements. *Proceedings of 2000 Military Sensing Symposia, Speciality Group on Passive Sensors, Ann Arbor, Michigan, August 2000*, 1, 169–181,
- Quegan, S. (1994). A unified algorithm for phase and cross-talk calibration of polarimetric data-theory and observations. *IEEE Trans.*, GRS-32, 89–99.
- Raney, R. K. (2006). Dual-polarized SAR and Stokes parameters. *IEEE Geoscience and Remote Sensing Letters*, 3 (3), 317–319.
- Raney, R. K. (2007). Hybrid-polarity SAR architecture. *IEEE Transactions on Geoscience and Remote Sensing*, 45 (11), 3397–3404.
- Reigber, A. and Moreira, A. (2000). First demonstration of airborne SAR tomography using multi-baseline L-band data. *IEEE Transactions on Geoscience and Remote Sensing*, 38/5, 2142–2152.
- Reigber, A., Papathanassiou, K. P., Cloude, S. R. and Moreira, A. (2001). SAR tomography and interferometry for the remote sensing of forested terrain. *Frequenz*, 55, 119–123.
- Roman, P. (1959a). Generalized Stokes parameters for waves with arbitrary form. *Il Nuovo Cimento*, **13**, 2546–2554.
- Roman, P. (1959b). Decomposition of 3×3 matrices. *Proc. Phys. Soc.*, 74, 649–657.
- Rosen, J. (1995). *Symmetry in Science: An Introduction to the General Theory*. Springer.
- Rosenqvist, A., Shimada, M., Ito, N. and Watanabe, M. (2007). ALOS PALSAR: A pathfinder mission for global-scale monitoring of the environment. *IEEE Transactions on Geoscience and Remote Sensing*, GRS45(11), 3307–3316.
- Sagues, L., Lopez-Sanchez, J. M., Fortuny, J., Fabregas, X., Broquetas, A. and Sieber, A. J. (2000). Indoor experiments on polarimetric SAR interferometry. *IEEE Transactions on Geoscience and Remote Sensing*, GRS-38, 671–684.
- Sagues, L., Lopez-Sanchez, J. M., Fortuny, J., Fabregas, X., Broquestas, A. and Sieber, A. J. (2001). Polarimetric radar interferometry for improved mine detection and surface clutter rejection. *IEEE Trans. Geoscience and Remote Sensing*, GRS-39, 1271–1278.
- Sarabandi, K. (1992a). Derivation of phase statistics from the Mueller matrix. *Radio Science*, **27**, 553–560.
- Sarabandi, K., Pierce, L. E. and Ulaby, F.T. (1992b). Calibration of a polarimetric imaging SAR. *IEEE Transactions on Geoscience and Remote Sensing*, GRS-30 (3).
- Saxon, D. S. (1955). Tensor scattering matrix for the electromagnetic field. *Phys. Rev.*, **100**, 1771.
- Schmeider, R. (1969). Stokes algebra formalism. *Journal of the Optical Society* of America, **59**, 297–302.
- Schneider, R. Z., Papathanassiou, K. P., Hajnsek, I. and Moreira, A. (2006). Polarimetric and interferometric characterization of coherent scatterers in urban areas. *IEEE Transactions on Geoscience and Remote Sensing*, GRS-44, 971–984.

- Schou, J., Skriver, H., Nielsen, A. A. and Conradsen, K. (2003). CFAR edge detector for polarimetric SAR images. *IEEE Transactions on Geoscience* and Remote Sensing, **41** (1), 20–32.
- Schuler, D. L., Lee, J. S., Kasilingam, D. and Nesti, G. (2002). Surface roughness and slope measurements using polarimetric SAR data. *IEEE Transactions on Geoscience and Remote Sensing*, **40** (3), 687–698.
- Seymour, S. and Cumming, I. G. (1994). Maximum likelihood estimation for SAR interferometry. *Proceedings of IEEE Geoscience and Remote Sensing Symposium, IGARSS'94, Pasadena, USA.*
- Sharma, J. J., Hajnsek, I. and Papathanassiou, K. P. (2007). Vertical profile reconstruction with POLInSAR of a subpolar glacier. *Proceedings of IEEE Geoscience and Remote Sensing Symposium, IGARSS'07*, 1147–1150.
- Simon, R. (1987). Mueller matrices and depolarization criteria. *Journal of Modern Optics*, **34**, 569–575.
- Souyris, J. C., Imbo, P., Fjortoft, R., Mingot, S. and Lee, J. S. (2005). Compact polarimetry based on symmetry properties of geophysical media: the pi/4 mode. *IEEE Transactions on Geoscience and Remote Sensing*, **43** (3), 634–646.
- Stebler, O., Meier, E. and Nueesch, D. (2002). Multi-baseline polarimetric SAR interferometry: first experimental spaceborne and airborne results. *ISPRS Journal of Photogrammetry and Remote Sensing*, **56** (3).
- Stokes, G. G. (1852). On the composition and resolution of streams of polarized light from different sources. *Cambridge Philosophical Society*, **9**, 399.
- Strang, G. (2004). *Linear Algebra and its Applications*. Fourth edition. Brooks Cole.
- Tabb, M. and Carande, R. (2001). Robust inversion of vegetation structure parameters from low frequency polarimetric interferometric SAR. *Proceedings of IEEE International Geoscience and Remote Sensing Symposium (IGARSS 2001), Sydney, Australia, July 2001.*
- Tabb, M., Orrey, J., Flynn, T. and Carande, R. (2002a). Phase diversity: a decomposition for vegetation parameter estimation using polarimetric SAR interferometry. *Proceedings of the Fourth European Synthetic Aperture Radar Conference, EUSAR 2002*, 721–724.
- Tabb, M., Flynn, T. and Carande, R. (002b). Direct estimation of vegetation parameters from covariance data in POLINSAR. *Proceedings of IGARSS 2002, Toronto, Canada*, 1908–1910.
- Topp, G. C., Davis, J. L. and Annan, A. P. (1980). Electromagnetic determination of soil water content: measurements in coaxial transmission lines. *Water Resources Research*, **16**, 574–582.
- Touzi, R., Lopes, A., Bruniquel, J. and Vachon, P. W. (1999). Coherence estimation for SAR imagery. *IEEE Transactions on Geoscience and Remote Sensing*, 37/1, 135–149.
- Touzi, R. (2007). Target scattering decomposition in terms of roll-invariant target parameters. *IEEE Transactions on Geoscience and Remote Sensing*, **45** (1), 73–84.

- Tragl, K. (1990). Polarimetric radar backscattering from reciprocal random media. *IEEE Transactions on Geoscience and Remote Sensing*, 28, 856–864.
- Treuhaft, R. N., Madsen, S., Moghaddam, M. and van Zyl, J. J. (1996). Vegetation characteristics and underlying topography from interferometric data. *Radio Science*, **31**, 1449–1495.
- Treuhaft, R. N. and Cloude S. R. (1999). The structure of oriented vegetation from polarimetric interferometry. *IEEE Transactions on Geoscience and Remote Sensing*, **37**/**2** (5), 2620.
- Treuhaft, R. N. and Siqueria, P. (2000a). Vertical structure of vegetated land surfaces from interferometric and polarimetric radar. *Radio Science*, **35** (1), 141–177.
- Treuhaft, R. N., Law, B. E. and Asner, G. P. (2000b). Structural approaches to biomass monitoring with multibaseline, multifrequency, polarimetric interferometry. *Proceedings of the Third European SAR Conference* (*EUSAR*), *Munich, Germany*, 253–255.
- Treuhaft, R. N., Law, B. E. and Asner G. P. (2004). Forest attributes from radar interferometric structure and its fusion with optical temote sensing. *BioScience*, **56** (6), 561–571.
- Tsang, L., Kong, J. A. and Shin, R. T. (1985). *Theory of Microwave Remote Sensing*. Wiley Interscience.
- Ulaby, F. T., Moore, R. K. and Fung, A. K. (1982). *Microwave Remote Sensing: Active and Passive. Vol. II: Radar Remote Sensing and Surface Scattering and Emission Theory.* Addison–Wesley.
- Ulaby, F. T., Moore, R. K. and Fung, A. K. (1986). *Microwave Remote Sensing: Active and Passive. Vol. III: From Theory to Applications.* Artech House.
- Ulaby, F. T. and Elachi, C. (eds.) (1990). *Radar Polarimetry for Geoscience Applications*. Artech House, Norwood, MA.
- van Albada, M. P., van der Mark, M. B. and Lagendijk, A. (1988). Polarization effects in weak localisation of light. *Journal of Physics D*, **21** (105), 28–31.
- van de Hulst, H. C. (1981). Light Scattering by Small Particles. Dover Press.
- van der Mee, C. V. M. and Hovenier, J.W. (1992). Structure of matrices transforming Stokes parameters. J. Math. Phys., **33** (10), 3574–3584.
- van der Mee, C. V. M. (1993). An eigenvalue criterion for matrices transforming Stokes parameters. J. Math. Phys., **34**, 5072–5088.
- van Zyl, J. J., Zebker, H. A. and Elachi, C. (1987). Imaging radar polarization signatures: theory and observations. *Radio Science*, **22**, 529–543.
- van Zyl, J. J. (1989). Unsupervised classification of scattering behaviour using radar polarimetry data. *IEEE Transactions on Geoscience and Remote Sensing*, GE-27(1), 36–45.
- van Zyl, J. J. (1990). Calibration of polarimetric tadar images using only image parameters and trihedral corner reflectors. *IEEE Transactions on Geoscience and Remote Sensing*, GE-28, 337–348.
- Wanielik, G. amd Stock, D. J. R. (1992). A proposed polarimetric CFARdetector and analysis of its operation. In Boerner, W. M et al. (eds.), *Direct and Inverse Methods in Radar Polarimetry*, 2, 999–1010. Kluwer Academic Publishers.

- Wiener, N. (1930). Generalized harmonic analysis. Acta Mathematica, 55, 118–258.
- Woodhouse, I. H. and Turner, D. (2002). On the visualization of polarimetric response. *International Journal of Remote Sensing*, **24** (6), 1377–1384.
- Woodhouse, I. H. (2006). Predicting backscatter-biomass and height-biomass trends using a macroecology model. *IEEE Transactions on Geoscience and Remote Sensing*, GRS-44, 871–877.
- Wright, P. A., Quegan, S., Wheadon, N. S. and David Hall, C. (2003). Faraday rotation effects on L-band spaceborne SAR data. *IEEE Transactions on Geoscience and Remote Sensing*, GRS-41 (12), 2735–2744.
- Yamada, H., Yamaguchi, Y., Rodriguez, E., Kim, Y. and Boerner, W. M. (2001). Polarimetric SAR interferometry for forest canopy analysis by using the super-resolution method. *IEICE Transactions on Electronics*, E84-C (12), 1917–1924.
- Yamaguchi, Y., Moriyama, T., Ishido, M. and Yamada, H. (2005). Fourcomponent scattering model for polarimetric SAR image decomposition. *IEEE Transactions on Geoscience and Remote Sensing*, **43** (8), 1699–1706.
- Zebker, H. A. and Villasenor, J. (1992). Decorrelation in interferometric radar echoes. *IEEE Transactions on Geoscience and Remote Sensing*, **30** (5), 950–959.
- Zhou, Z. S. and Cloude, S. R. (2006). Application of polarization coherence tomography to GB-POLInSAR data. Proceedings of IEEE International Symposium on Geoscience and Remote Sensing, IGARSS06, Denver, Colorado, July 2006.

This page intentionally left blank

Index

along track interferometry (ATI), 219 ALOS-PALSAR, 396 alpha parameter and dihedral dielectric constant, 399 definition, 184 for Bragg scattering, 128 for dihedral scattering, 130 mean alpha, 98 anisotropy scattering parameter A, 97

baseline components, 213
baseline decorrelation, 224
biaxial material, 13
bidirectional reflectance distribution function (BRDF), 139
birefringence

for circular polarisations, 24
general definition, 14

bivariate Gaussian distribution, 429
Bragg surface scattering, 126
Brewster angle, 122

C2 symmetry, 11 calibration of POLInSAR systems, 361 of POLSAR systems, 351 polarimetric, 350 Quegan algorithm, 352 Cameron decomposition, 182 Cartan matrix, 417 Cartan sub-algebra, 416 application to coherency matrix, 92 Chandrasekhar decomposition, 191 chiral media admittance parameters, 23 D and L-rotatory materials, 24 spatial dispersion, 23 specific rotatory power of, 25 wave equation for, 23 Cloude-Pottier decomposition, 192 coefficient of variation (CV), 427 coherence for two-layer surface/volume problems, 270 for circular polarisation, 136 for exponential profile, 231

for semi-infinite random volume, 258 SNR decorrelation, 221 geometric decorrelation, 224 image co-registration errors, 223 interferometric, 220 polarimetric, 73 polarimetric interferometric, 235 temporal decorrelation, 222 coherence loci definition, 252 for IWCM model, 280 for oriented volume scattering, 263 for OVOG model, 282 for OVUG model, 283 for random volume scattering, 257 for RVOG, 274 for RVUG model, 283 for surface scattering, 255 coherence optimization estimation bias, 250 constrained, 243 for oriented volume scattering, 261 for random volume scattering, 255 for surface scattering, 254 SVD interpretation, 242 unconstrained, 241 coherence region, 245 coherence tomography condition number for multi-baselines, 338 dual baseline, 325 multi-baseline reconstruction, 335 single baseline, 323 single baseline condition number, 332 temporal decorrelation and SNR effects. 334 coherency matrix and CFAR detection, 194 and contrast optimization, 193 for forward scattering, 165 for general scattering, 86 for waves, 73 propagation effects, 202 compact polarimetry, 354 and Bragg scattering, 176 and Rayleigh scattering by spheroids, 176 $\pi/4$ mode. 356 compact POLInSAR, 362 condition number of a matrix, 409

contrast optimization, 193 COPOL nulls, 57 Cramer–Rao bounds, 431 critical baseline, 215 cross-polarisation, 8

decomposition eigenvalue decomposition of [J], 74 eigenvalue decomposition of [T], 87 of interferometric coherence, 233 of N-dimensional coherency matrices, 92 of the Stokes vector, 78 of the wave coherency matrix, 76 propagation distortions, 205 roll invariance, 179 the point reduction theorem, 185 degree of polarisation, 76 depolarisation definition, 71 state vector, 92 depolarisers azimuthal depolarisation, 96 characterisation of, 110 isotropic, 91 reflection depolarisation, 97 dielectric constant of soil. 120 Polder-Van Santen/de Loor formula, 171 soil moisture, 121 soil salinity, 120 dielectric tensor, 12 differential interferometry, 217 dihedral scattering from dielectrics, 124 from metal structures, 54 Dirac matrices, 413 directional hemispherical reflectivity (DHR), 139 discrete dipole approximation (DDA), 152 DLR E-SAR, 392 Dynkin diagrams, 418

effective length of an antenna, 51 effective number of looks (ENL), 427 eigenpropagation states, 14 entropy backscattering, 97 general scattering entropy, 108 of a wave, 77 entropy-alpha decomposition, 194 entropy-alpha diagram for backscatter, 99 for bistatic scattering, 109 for compact polarimetry, 101 for dual polarisation, 101

Faraday rotation basic properties, 21 estimation from data, 206 observations from satellites, 397 vectorised form, 205 field of values of a matrix, 243 flat earth phase component, 210 flat earth phase removal, 216 Foldy Lax Equations, 33 forest propagation extinction models, 174 Fourier-Legendre coherence basis functions, 228 Fourier-Legendre series, 225 Freeman-Durdan decomposition, 197 Freeman-eigenvalue hybrid decomposition, 198 Fresnel equations, 117 Fry-Kattawar relations, 82

Gell–Mann matrices, 414 Graves power matrix, 48 gyrotropic media, 18

height estimation structure free algorithm, 306 using RVOG, 309 Hermitian matrices, 404 homomorphisms of Lie groups, 420 Huygens source, 8 Huynen decomposition, 192 Huynen parameters, 63

InSAR, 345 integral equation model (IEM), 138 interferometric water cloud model (IWCM), 278 interferometry blind angles, 214 memory line, 212 vertical wavenumber, 213 π -height, 214 ionosphere basic properties, 18 cyclotron frequency, 19 extraordinary wave, 20 ordinary wave, 20 plasma frequency, 18 isomorphism, 41

Jones calculus definition, 25 diattenuator, 26 homogeneous propagation channel, 28 inhomogeneous propagation channel, 29 N matrix, 25 retarder, 27

Kennaugh matrix, 83 Killing form, 416 Krogager decomposition, 180

Lagrange multipliers and coherence optimization, 241 and Rayleigh quotient, 407 and the [S] matrix, 48 Lambertian surface, 139 leaf-area-index (LAI), 173 Lee filter, 348, 427 Legendre coherence model first order, 296 second order, 301 Lie algebra, 412 line fit algorithm for two states, 287 total least squares (TLS), 290 log-likelihood function, 433 Lorentz force equation, 20 Lorentz spin matrix, 43 Lorentz transformation, 43 and [S] matrix, 67 connection to special relativity, 66 conservation of zero wave entropy, 82 homomorphism, 64 Lorentz boost, 66 Ludwig wave co-ordinates, 6

Malus's law, 64 matrix exponential function, 15 Maxwell's equations boundary conditions, 115 coherent surface scattering, 129 differential form, 3 dipole radiation, 5 duality transformation, 6 Green's function, 4 Helmholtz equation, 4 inhomogeneous plane wave, 119 physical optics approximation, 123 TE/TM Waves, 116 TEM waves, 5 vector wave equation, 3 Mie scattering, 152 Minkowski metric, 43 Mishchenko decomposition, 164 Monte Carlo polarisation simulations, 432 Mueller matrix backscatter form. 83 definition. 79 filtering, 90 for isotropic depolariser, 84 formal connection to [T], 113 main properties, 80 mapping to [T], 88 pure matrices, 80 reciprocity theorem for backscatter, 85 sum of pure matrices, 86 test for a pure matrix, 90 the phase function, 153 multivariate normal distribution, 433

N matrix, 14 noise estimation from crosspolar channels, 189 in coherence tomography, 333 in decomposition theorems, 189 in interferometry, 221 norm of a matrix, 409 nullspace and decomposition theorems, 191 definition, 406

orthogonal scattering mechanisms definition, 186 in natural media, 187 OVOG model, 281 OVUG model, 283

Pancharatnam phase, 37 Pauli spin matrices, 16 penetration depth, 142 permanent scatterers (PS), 224 phase bias removal, 286 plane of polarisation, 11 Poincaré sphere, 42 polar decomposition of a matrix, 27 polarimetric interferometry change of wave basis, 237 forming vector interferograms, 235 optimum baseline, 318 phase normalization, 235 SVD and Schur decompositions, 248 polarisation C-lines. 37 conjugate semi-diameters, 36 definition of sense, 19 equation of ellipse, 35 left and right circular, 19 L-lines, 37

polarisation coherence tomography (PCT), 329 polarisation fork, 62 polarisation frame or basis, 11 polarisation signatures, 60 polarization of matter, 10 POLINSAR, 360 POLSAR, 347 Poynting vector, 37 principal material axes, 13

range spectral filtering, 212 Rayleigh quotient and optimum transmittance, 29 definition, 405 Rayleigh scattering, 145 and the entropy-alpha plane, 159 bistatic scattering, 160 by a chiral spheroid, 150 by a cloud of chiral spheroids, 158 by a cloud of spheroids, 156 by a small spheroid, 148 particle anisotropy A_p , 148 spheroidal particle functions, 149 reciprocity theorem, 49 refractive index, 11 retarder, 18 RVOG model, 271 minimum coherence, 275

scattering amplitude matrix backscatter theorem, 50 bisectrix, 46 BSA co-ordinate system, 50 definition of [S], 47 FSA co-ordinate system, 50 scattering angle, 46 scattering plane, 46 singular matrices, 64 transformation invariants, 58 scattering matrix group, 49 scattering mechanism, 69 scattering sphere application to particle scattering, 154 definition, 108 scattering symmetries backscatter azimuthal symmetry, 95 backscatter reflection symmetry, 94 backscatter rotation symmetry, 95 bistatic symmetries, 104 degree of symmetry, 183 symmetric scatterers, 181 scattering vector lexicographic order, 68 Pauli expansion, 68 Schur decomposition and coherence, 247 Schur's theorem, 406 Schwarz inequality, 428 similarity transformations, 405 SINC coherence model, 229 singular value decomposition and coherence optimization, 242 and transformation of the [S] matrix, 55 of the scattering amplitude matrix, 47 SVD, 407 skin depth, 119 small perturbation model (SPM), 125 Snell's law, 116 speckle filtering, 427 sphere-diplane-helix (SDH) decomposition, 181 spin matrix, 40 Stokes criterion, 84 Stokes reflection matrix, 83 Stokes vector, 40 connection to w, 114 definition, 43 for three-dimensional Waves, 45 surface slope azimuth slope, 131 range slope, 131 surface topography estimation for the OVOG model, 293 for the OVUG model, 294 for the RVOG model, 287 surface-to-volume ratio estimation from polarimetry, 200

RVOG estimation, 316 synthetic aperture radar (SAR), 341

T-matrix method, 152 trace of a matrix, 403

Ulaby scattering model, 170 uniaxial material, 13 unitary matrix definition, 404 unitary transformations and double angles, 41 Cartan sub-algebra, 74 compound form for SU(2), 39 congruent, 56 homomorphic, 41 special unitary, 12 unitary reduction operator, 91

vector radiative transfer (VRT) cyclical components, 164 ladder terms, 164 vegetation bias, 229 vertical structure function definition, 225 estimation using CT, 323

water cloud model (WCM), 168 wave dichotomy, 75 Wishart distribution, 430

X-Bragg model and entropy-alpha plane, 136 definition, 134 effect of surface slope, 136 moisture parameter M, 135 roughness parameter R, 135 XPOL nulls, 57

Yamaguchi decomposition, 198