HydroPol
Mathematical Documentation

****

Invariant Imbedding Theory
for the
Vector Radiative Transfer Equation

Curtis D. Mobley
Sequoia Scientific, Inc.
2700 Richards Road, Suite 107
Bellevue, WA 98005

Version 1.0

February 12, 2016

.

Executive Summary

This report develops in detail the mathematics needed to solve the vector (polarized) radiative transfer equation (VRTE) using invariant imbedding theory. The scalar (unpolarized) version of invariant imbedding theory was developed for the oceanographic setting in the text *Light and Water* (Mobley, 1994). This report parallels that work and extends the scalar equations to the vector level. The presentation of the vector-level equations is complete, including generic internal sources and wind-blown sea surfaces with multiple scattering by gravity-capillary waves. Issues of computer array storage are addressed, and flow charts of the computations are given. The vector-level invariant imbedding theory as developed here is new.

Chapter 1 formulates the vector radiative transfer problem for prediction of the time-averaged propagation of polarized light in a plane-parallel ocean. Chapter 2 shows how the continuous-variable equations of Chapter 1 are discretized for solution on a computer.

Chapter 3 shows how the reflectance and transmittance properties of wind-blown sea surfaces can be computed using Monte Carlo simulation. The Fresnel reflectance and transmission formulas for polarized light are first presented for a level sea surface. Wind-blown sea surfaces are then discussed in detail. The generation of random realizations of sea surfaces is reviewed for techniques using either the Cox-Munk wave-slope model, or using wave energy spectra and fast Fourier transforms to generate random surfaces for fully developed seas. Monte Carlo ray tracing is then developed in great detail for arbitrary sea surfaces. This algorithm goes beyond any previous work and enables the computation of polarized reflectance and transmission by fully developed seas, including multiple scattering of light by the surface waves.

Chapter 4 then develops the invariant imbedding solution of the in-water VRTE. The VRTE is partitioned into separate sets of equations for upwelling and downwelling light, which is a key feature of invariant imbedding theory. These equations are then Fourier decomposed in the azimuthal direction. This leads first to local interaction equations, which govern how infinitesimal slabs of water reflect and transmit light, and then to global interaction equations, which govern how finitely thick slabs of water reflect and transmit light. Upward and downward sets of differential equations are developed for the operators occurring in the global interaction equations. Given the absorbing and scattering properties of the water, solution of these equations gives a "bare slab" or "boundaryless" solution of the VRTE within the water column.

Chapter 5 then shows how to incorporate the boundary conditions at the sea surface and sea bottom into the bare slab solution for the water column. Those boundary conditions define specific environmental conditions such as the incident sun and sky radiance, surface wave state, and bottom reflectance. Combining the boundary conditions with the bare slab internal solution gives the solution of the VRTE for the full system of sea surface, water column, and bottom boundary for a specific set of inputs. Adding the boundaries to the water body leads to vector-level invariant imbedding relations, imbed rules, and union rules corresponding to–but much more complex than–those found in the scalar theory. The final output of these equations is the Stokes vector within and leaving the water body.

Chapter 6 gives a flow chart of the entire computation process. Two appendices complete the report. The first supplies for reference the formulas for Fourier decomposition of functions of one

or two discrete variables. The second gives the algorithm for ray tracing in three-dimensional space as a photon interacts with a random sea surface.

In the interest of completeness, this report also contains much material that is considered "well known." However, in reality, this information is spread among many publications and is confused by different notations, conventions, and terminologies. It is hoped that parts of this report may therefore also serve as a tutorial on the basics of polarized radiative transfer for those who do not wish to delve into the mathematics of invariant imbedding theory. There is emphasis throughout on the physical interpretation of the often complicated equations. These ancillary discussions are one of the luxuries of a technical report versus a journal article.

The widely used HydroLight software uses the scalar versions of the equations to solve the scalar radiative transfer equation. Corresponding software, called HydroPol (HydroLight-Polarized), is under development to implement the vector-level solution techniques developed here. The HydroPol source code is documented via references to the equations in this report.

This report is available online as Mobley(2014) in the references section of the *Ocean Optics Web Book* at http://www.oceanopticsbook.info/view/references/publications

Dedication

Rudolph W. Preisendorfer (1927-1986) was one of the founding fathers of hydrologic optics. He was a pencil-and-paper mathematician whose greatest contribution to science was the development of invariant imbedding theory for solution of the scalar (unpolarized) radiative transfer equation in the oceanographic setting. That work and much more is presented in his six-volume treatise *Hydrologic Optics* (Preisendorfer, 1976). The mathematics he developed for the scalar equation is the core of the HydroLight software, which I began developing with him in the years before his untimely death. Invariant imbedding theory has proved to be robust, accurate, and extremely fast compared to many other techniques for solving the scalar radiative transfer equation. Preisendorfer however said almost nothing about polarization and, to my knowledge, neither he nor anyone else has extended invariant imbedding theory to the vector (polarized) radiative transfer equation in the way it is done here for oceanic computations.

It was my privilege to work with Rudy during my postdoctoral and early career years (1977-1986). If I have accomplished anything in the present work, it is because he showed the way. I humbly dedicate this extension of his work to his memory.

Acknowledgment

This work was supported by NASA Contract NNH12CD06C titled *Radiative Transfer Modeling for Improved Ocean Color Remote Sensing*. This report constitutes part of the final report on that contract.

★ ★ ★

"Great God! this is an awful place and terrible enough for us to have laboured to it.... Well, it is something to have got here,.... Now for the run home and a desperate struggle. I wonder if we can do it."

–R. F. Scott upon reaching the South Pole, Jan 17, 1912

"Great God! this is an awful derivation and terrible enough for me to have labored to it.... Well, it is something to have got here,.... Now for the programming a desperate struggle to debug. I wonder if I can do it."

–C. D. Mobley upon finishing the HydroPol algorithm, 102 years later

Caveat to the Reader

This document shows my development of invariant imbedding theory for the vector radiative transfer equation in a plane parallel ocean. When I started this work, I thought it would be a straightforward extension of the mathematics for the scalar case as developed in *Hydrologic Optics* and restated in *Light and Water* (Mobley, 1994). I was in for a rude awakening. A number of unforeseen complications arose because of loss of certain symmetries in going from the scalar to the vector theory. The mathematics thus turned out to be unexpectedly messy and the development was very time consuming.

I have done my best to check and check again that my derivations are correct. However, there remains much room for error, both in the formulation of certain parts of the theory (in particular in the way the global interaction principles are formulated) and as simple typos. The reader is therefore cautioned that this document is just "Version 1.0" of the HydroPol solution algorithm, and almost no equation seen here is absolutely guaranteed to be correct. I have added footnotes at several places where something unexpected occurred in the development, or where a bit of philosophical explanation is needed. I truly wish Rudy Preisendorfer were still here to check the equations presented below, especially those in Chapters 4 and 5 .

Two topics have been omitted from this first version of these notes: the detailed formulation of internal source terms for fluorescence and Raman scatter by water, and the formulation of the polarized reflectance properties of an infinitely deep layer of water. The first is needed for the inclusion of specific types of inelastic scatter in the solution, and the second enables the simulation of an infinitely deep water body below the greatest depth of interest. Those matters will be considered after the code is fully debugged for a source-free, finite-depth water body. Thus these notes will be revised as the HydroPol software is developed.

The only way to know if I have done everything correctly is if these equations, when programmed into the HydroPol software, lead to computed underwater Stokes vectors that agree with measurements and with independently developed Monte Carlo codes for prediction of underwater polarized light fields. That work is under way but is far from complete at the time of this writing. Meanwhile, if you think you found an error, you probably did. I welcome your comments and corrections.

Contents

List of Figures

CHAPTER 1

Formulation of the Vector Radiative Transfer Problem

These notes presume a basic knowledge of radiometry, optical oceanography, and polarization as given in *Light and Water* (Mobley, 1994), Kattawar and Adams (1989), Kattawar (1994), and the introductory parts of Mischenko et al. (2002). The mathematics developed here to solve the vector (or polarized) radiative transfer equation (VRTE) parallels that for the scalar (unpolarized) radiative transfer equation (SRTE). The development and solution of the SRTE as used in the unpolarized HydroLight software are given in detail in *Light and Water*, which guides the development of the corresponding vector equations. Occasional reference will therefore be made to the scalar equations in *Light and Water*, with those scalar equations being prefaced here by an "L&W", e.g. Eq. (L&W 8.74). The mathematical development in this report is extremely detailed compared to a journal article describing the same material. This level of detail is necessary because the associated computer code must implement these equations down to the last subscript and array index. The HydroPol source code is documented with comments linking to the corresponding equations of this report.

There are many numerical methods for solving the VRTE. Many of the pioneering papers are collected in Kattawar (1991). Some of these methods seem to have been published and used for studies in a few papers, but never found widespread usage. Other techniques–Monte Carlo simulation in particular–have found widespread application by many users. The solution technique used in HydroPol is an extension of the invariant imbedding algorithm used for the SRTE in the widely used HydroLight software. Invariant imbedding is mathematically complicated, but its computational efficiency and accuracy have been well established by over two decades of HydroLight usage. Invariant imbedding has several advantages over other solution techniques, the most important of which are

1. Radiances are computed in all directions simultaneously (some techniques compute the radiance in only one direction per run of the computer code).
2. All orders of multiple scattering are included in the calculations (some techniques include only the first few orders of multiple scattering).
3. Computer run time is linearly proportional to the optical depth to which the VRTE is solved (run time for some techniques increases exponentially with optical depth).

As will been seen, the HydroPol invariant imbedding algorithm involves several steps:

1. The VRTE and associated boundary conditions are directionally discretized by integrating continuous functions of direction over finite solid angles. The numerically computed Stokes vectors are then exact values averaged over the finite solid angles.

2. The VRTE is rewritten as a pair of equations for upward and downward directions.

3. The VRTE is Fourier decomposed in azimuthal direction, which allows the solution in the interior of the water body to be computed as a sequence of independent "small" problems, rather than as one "large" problem.

4. These upward and downward equations are reformulated in terms of reflectance and transmittance functions for finitely thick layers of water. Those reflectance and transmittance functions depend only in the inherent optical properties of the water body and can be computed independently of the particular boundary conditions (i.e., incident lighting and bottom reflectance) applied to the water body.

5. The surface and bottom boundary conditions are then applied to the generic interior solution to obtain the solution of the VRTE for specific conditions of incident lighting, bottom reflectance, and internal sources.

6. Finally, the physical Stokes vectors are then reconstituted from the Fourier amplitudes.

Invariant imbedding as outlined above was developed by Preisendorfer for solution of the SRTE in a series of papers culminating in his *Hydrologic Optics* treatise (Preisendorfer, 1976). The algorithm as applied to the SRTE is described in detail in Chapters 7 and 8 of *Light and Water*. The extension of the scalar invariant imbedding equations to the VRTE as given in this report is apparently new.

1.1 Terminology and Notation

The state of polarization of a light field is specified by the four-component *Stokes vector*, whose elements are related to the complex amplitudes of the electric field vector $\boldsymbol{E}$ resolved into directions that are parallel ($E_\parallel$) and perpendicular ($E_\perp$) to a conveniently chosen reference plane. However, there are two versions of the Stokes vector seen in the literature, and these two versions have different units and refer to different physical quantities. The *coherent* Stokes vector describes a quasi-monochromatic plane wave propagating in one exact direction, and the vector components have units of power per unit area (i.e., irradiance) on a small surface element perpendicular to the direction of propagation. To be specific, the coherent Stokes vector is defined as (Zhai et al., 2012)

$$\underline{S} = \begin{bmatrix} I \\ Q \\ U \\ V \end{bmatrix} = \sqrt{\frac{\epsilon}{\mu_m}} \begin{bmatrix} E_\parallel E_\parallel^* + E_\perp E_\perp^* \\ E_\parallel E_\parallel^* - E_\perp E_\perp^* \\ E_\parallel E_\perp^* + E_\perp E_\parallel^* \\ i(E_\parallel E_\perp^* - E_\perp E_\parallel^*) \end{bmatrix} . \tag{1.1}$$

Here ϵ is the electric permittivity of the medium, which has units of Farad/m^2 or SI units of $A^2\,s^4\,kg^{-1}\,m^{-3}$; and μ_m is the magnetic permeability of the medium with units of Henry/m or $kg\,m\,s^{-2}\,A^{-2}$. Electric fields have units of Newton/coulomb or $kg\,m\,s^{-3}\,A^{-1}$. Thus the elements of the Stokes vector have units of $kg\,s^{-3}$ or Watt/m^2. E^* denotes complex conjugate, hence the components of the Stokes vector are real numbers.

The *diffuse* Stokes vector is defined as in Eq. (1.1) but describes light propagating in a small set of directions surrounding a particular direction and has units of power per unit area per unit solid angle (i.e., radiance). It is the diffuse Stokes vector that appears in the radiative transfer equation as developed here. The differences in coherent and diffuse Stokes vectors are rigorously presented in Mischenko (2008).

Some authors (e.g. Bohren and Clothiaux, 2006) omit the $\sqrt{\epsilon/\mu_m}$ factor in Eq. (1.1) because they are interested only in relative values such as the degree of polarization or other ratios of the components of the Stokes vector, not absolute magnitudes, but this omission is both confusing and physically incorrect. *Units and magnitudes matter!* The different units of coherent and diffuse Stokes vectors, and whether or not the $\sqrt{\epsilon/\mu_m}$ factor is included in the definition of the Stokes vector, have subtle but very important consequences in how light propagation across a dielectric interface such as the air-water surface is formulated, as will be seen below. The recent paper by Zhai et al. (2012) gives a definitive discussion of these matters.

HydroPol uses SI units for all quantities except wavelength, which is in nanometers as is customary in optical oceanography. Thus spectral radiance has units of $W\,m^{-2}\,sr^{-1}\,nm^{-1}$ and spectral irradiances have units of $W\,m^{-2}\,nm^{-1}$. The mathematics needed to solve the VRTE using invariant imbedding theory becomes quite complicated, e.g., matrices whose elements are matrices, all of which must be turned into a computer program. Precise terminology and notation are required to keep everything straight. Table 1.1 shows the notation used.

Notation	**Examples**	**Usage**
primed variables	$\theta', \phi', \boldsymbol{\xi}'$	incident or unscattered directions
unprimed variables	$\theta, \phi, \boldsymbol{\xi}$	final or scattered directions
bold face	$\mathbf{p}, \boldsymbol{\xi}$	points in space or unit vectors defining directions in 3D space
underline	$\underline{S}$, $\underline{P}$	4×1 Stokes vectors $\underline{S}$ or 4×4 matrices whose elements are numbers
double underline	$\underline{\underline{R}}$, $\underline{\underline{\tau}}$	composite matrices whose elements are themselves matrices or vectors
tilde	$\underline{\tilde{M}}$, $\tilde{P}_{i,j}$	matrices or their elements that have been "reduced" by factoring out the (1,1) element of the corresponding matrix $\tilde{P}$, so that $\tilde{P}_{i,j} = P_{i,j}/P_{1,1}$; hence $\tilde{P}_{1,1} = 1$.
caret or hat	$\hat{\underline{\tau}}$, $\hat{P}_{i,j}$	Fourier amplitudes of the respective (unhatted) physical quantity

Table 1.1: Notation.

1.2 Geometry

We next choose coordinate systems and show in detail how to resolve Stokes vectors in these coordinate systems as needed for scattering calculations.

1.2.1 Coordinate Systems

Figure 1.1 shows the coordinate systems used in HydroPol. Depth and direction are defined in a 3D Cartesian coordinate system with depth measured positive downward from 0 at the mean sea surface. Polar angle θ is defined from 0 in the $+\mathbf{z}$ or downwelling direction to π in the $-\mathbf{z}$ upwelling direction. $+\mathbf{x}$ points in the downwind direction; thus $\pm\mathbf{y}$ is the cross-wind direction. As will be seen in §5.1, using a wind-centered coordinate system makes it easier to model a random sea surface with different along-wind and cross-wind slope statistics. Azimuthal angle ϕ is measured counterclockwise from $+\mathbf{x}$ when looking in the $-\mathbf{z}$ direction. The sun is placed at an azimuthal angle of ϕ_{sun} relative $+\mathbf{x}$. If the sun is placed in the $+\mathbf{x}$ azimuthal direction at $\phi_{\text{sun}} = 0$, unscattered photons from the sun are then traveling in the $-\mathbf{x}$ direction at $\phi = 180$ deg.

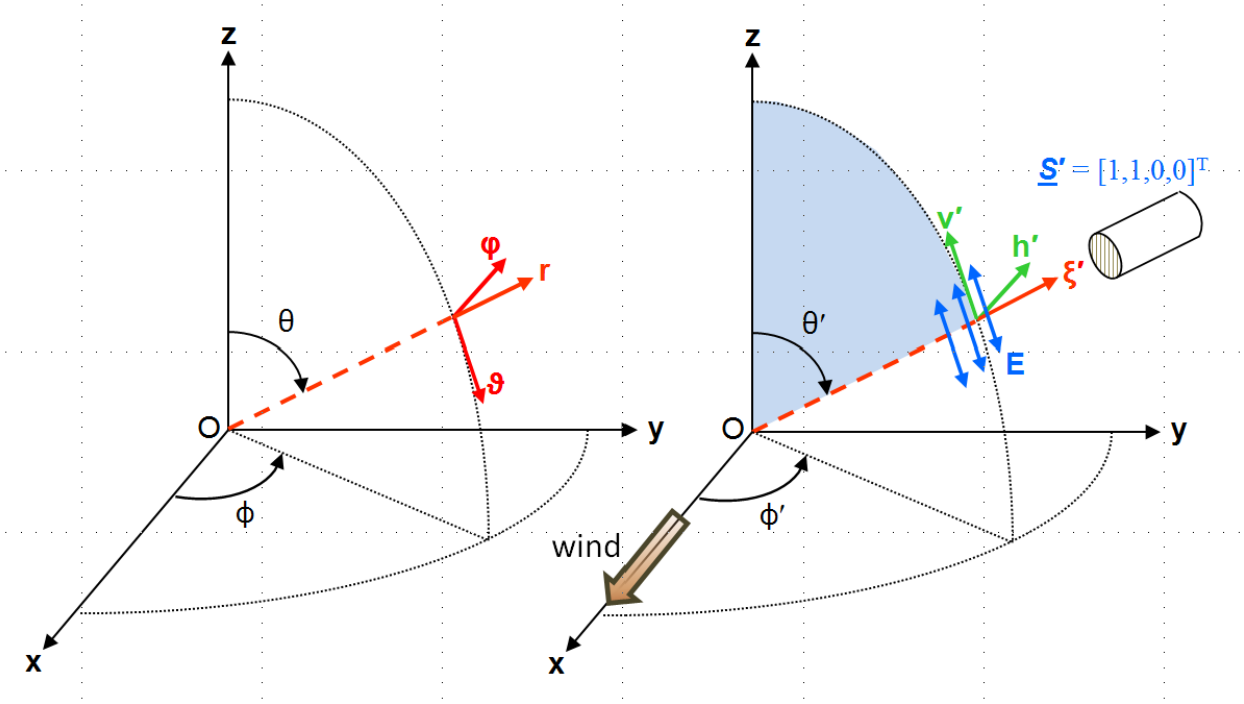

Figure 1.1: HydroPol coordinate systems. The left panel shows the **x**-**y**-**z** Cartesian system, which is fixed in space and used to define the spherical (r, θ, ϕ) coordinate system. The right panel shows an incident photon direction $\boldsymbol{\xi}'$ whose Stokes vector is $\underline{S}'$. The $\mathbf{h}'$-$\mathbf{v}'$-$\boldsymbol{\xi}'$ system is a local system (changing with θ', ϕ') for defining photon direction and resolving Stokes vectors with respect to directions that are perpendicular and parallel to the incident meridian plane, part of which is shaded in blue. In HydroPol, this figure is turned "upside down," with the **x**-**y** plane being the mean sea surface, **x** pointing downwind, and **z** pointing downward into the water.

The unit vectors in the directions of increasing (r, θ, ϕ) are

$$\mathbf{r} = \sin\theta\cos\phi\,\mathbf{x} + \sin\theta\sin\phi\,\mathbf{y} + \cos\theta\,\mathbf{z}\,, \tag{1.2a}$$

$$\boldsymbol{\vartheta} = \cos\theta\cos\phi\,\mathbf{x} + \cos\theta\sin\phi\,\mathbf{y} - \sin\theta\,\mathbf{z}\,, \tag{1.2b}$$

$$\boldsymbol{\varphi} = -\sin\phi\,\mathbf{x} + \cos\phi\,\mathbf{y}\,. \tag{1.2c}$$

Let $\boldsymbol{\xi}$ denote a unit vector pointing in the direction of light propagation, as given by angles (θ, ϕ). (In Fig. 1.1, $\boldsymbol{\xi}' = \mathbf{r}$.) The components of $\boldsymbol{\xi}$ are given by

$$\boldsymbol{\xi} = \xi_x\mathbf{x} + \xi_\mathbf{y}\mathbf{y} + \xi_\mathbf{z}\mathbf{z} = (\xi_\mathbf{x}, \xi_\mathbf{y}, \xi_\mathbf{y}) \tag{1.3a}$$

where

$$\xi_x = \sin\theta\cos\phi\,, \tag{1.3b}$$

$$\xi_y = \sin\theta\sin\phi\,, \tag{1.3c}$$

$$\xi_z = \cos\theta\,. \tag{1.3d}$$

In radiative transfer theory, direction (θ, ϕ) always refers to the direction of photon travel. Thus light traveling straight down is traveling in the $+\mathbf{z}$ or $\theta = 0$ direction. To detect this radiance, an instrument is pointed in the $-\mathbf{z}$ or $\theta = \pi$ direction, which is termed the zenith-viewing direction. Radiance traveling straight upward in the $-\mathbf{z}$ or $\theta = \pi$ direction is referred to as the nadir-viewing radiance since it is detected by an instrument pointing in the nadir $(\theta = 0)$ direction. Radiance is often more conveniently plotted as a function of the viewing direction $(\theta_v, \phi_v) = (\theta - \pi, \phi + \pi)$.

We will generally use a prime to denote an incident or unscattered direction, e.g, $\boldsymbol{\xi}'$ or (θ', ϕ'). Unprimed variables denote final or scattered directions, e.g. $\boldsymbol{\xi}$ or (θ, ϕ). However, in the development of the sea-surface reflectance and transmittance equations in Chapter 3, it will be more convenient to use subscripts i, r, and t for incident, reflected, and transmitted (refracted) directions, respectively, e.g. $\boldsymbol{\xi}_i, \boldsymbol{\xi}_r$, and $\boldsymbol{\xi}_t$.

We will have frequent need to define unit vectors in directions perpendicular and parallel to a given plane (which can be a meridian plane, a scattering plane in the water volume, or the plane of incident, reflected, and transmitted light for light incident onto an air-water or bottom surface, all of which are defined below). HydroPol uses the following conventions for defining these unit vectors. It is common in the literature to let $\mathbf{p}$ denote a vector parallel to a plane, and $\mathbf{s}$ denote a vector perpendicular (*senkrecht* in German) to a plane. The perpendicular vector $\mathbf{s}$ is chosen to be in the direction given by the vector cross product of the incident direction crossed with the final direction. The parallel vector $\mathbf{p}$ is then defined as the direction of propagation cross the perpendicular direction. Thus the perpendicular cross parallel directions give the direction of light propagation: $\mathbf{s} \times \mathbf{p} = \boldsymbol{\xi}$, where $\times$ denotes the vector cross product.

For an incident direction $\boldsymbol{\xi}'$ and the associated Stokes vector $\underline{S}'$ specified in the incident meridian plane, the first vector is taken to be $\mathbf{z}$ and the second is the direction of propagation. Thus $\mathbf{h}' \equiv \mathbf{z} \times \boldsymbol{\xi}'/|\mathbf{z} \times \boldsymbol{\xi}'|$ as seen in the right panel of Fig. 1.1. In this case $|\mathbf{z} \times \boldsymbol{\xi}'| = \sin\theta'$, and $\mathbf{h}' = \boldsymbol{\varphi}'$. Similarly, $\mathbf{v}' \equiv \boldsymbol{\xi}' \times \mathbf{h}' = -\boldsymbol{\vartheta}'$, and $\mathbf{h}' \times \mathbf{v}' = \boldsymbol{\xi}'$. In our geometry, meridian planes are perpendicular to the mean sea surface. The $\mathbf{h}'$ vector as just defined is parallel to the mean sea surface and therefore is often referred to as the "horizontal" direction; $\mathbf{v}'$ lies in a vertical plane and is correspondingly called the "vertical" direction. For a final direction $\boldsymbol{\xi}$ and its Stokes vector $\underline{S}$ in the final meridian plane, the first vector is the direction $\boldsymbol{\xi}$ and the second is $\mathbf{z}$.

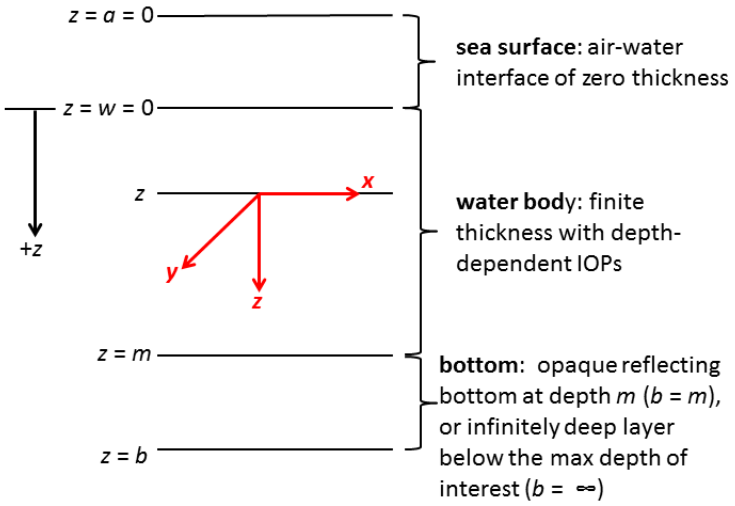

Figure 1.2: Partition of the entire water body into surface, water column, and bottom layers.

This vector cross product algorithm for specifying perpendicular and parallel directions will be convenient for sea surface reflectance and transmission calculations in which light can propagate from one tilted wave facet to another without reference to meridian planes, except for the incident and final directions when a photon enters or leaves the region of the sea surface.

The Q and U components of a Stokes vector describe linear polarization with the plane of polarization specified relative to a particular coordinate system. The I component is the total radiance, and V describes circular polarization; these quantities do not depend on the coordinate system and are invariant under a rotation of the coordinate system. The blue arrows in Fig. 1.1 represent the plane of oscillation of the electric field vector $\boldsymbol{E}$ parallel to the meridian plane, i.e. for vertical plane polarization. The components of $\underline{S}' = [1, 1, 0, 0]^T$ as shown thus represent radiance of magnitude 1 $\mathrm{W\,m^{-2}\,sr^{-1}\,nm^{-1}}$ that is 100% vertically plane polarized.

For the development of invariant imbedding theory for the VRTE, it will be convenient to divide the entire water body into the sea surface, the water column, and the bottom, as shown in Fig. 1.2. Depth $z = a = 0$ refers to a location in the air just above the mean sea surface; depth $z = w = 0$ is in the water just below the mean sea surface. The sea surface is thus a "layer" of zero thickness, which physically represents a discontinuity in the index of refraction between the air above and the water below. This layer will be denoted $[a, w]$. The water column extends from depth $z = w = 0$ to some maximum depth of interest at $z = m$. This user-specified depth m is the maximum depth to which the VRTE will be solved. The water body can be divided into sublayers, e.g., $[w, z]$ is the water from the surface to any depth z, and $[z, m]$ is the water from depth z to the maximum depth of interest. The IOPs within the water body are generally functions of depth; the water-column layer notation $[w, z]$ and $[z, m]$ does not imply that the layers are homogeneous. A bottom boundary condition is applied at depth m. This bottom boundary condition can describe either an opaque reflecting bottom physically located at depth $z = m = b$, or it can describe the reflectance properties of an infinitely deep, homogeneous, source-free water body below depth m, in which case $b = \infty$ and the bottom layer of water is denoted $[m, \infty]$. The entire water column is the union of these surface, water, and bottom layers: $[a, b] = [a, w] \cup [w, m] \cup [m, b]$.

1.2.2 Scattering Geometry

In HydroPol the elements of input and output Stokes vectors are defined relative to meridian planes, as described above. However, scattering from an incident direction $\boldsymbol{\xi}'$ to a final direction $\boldsymbol{\xi}$ is defined in terms of the included scattering angle ψ and the *scattering plane*, as illustrated in Fig. 1.3. Using Eq. (1.3) to express the incident direction $\boldsymbol{\xi}'$ and scattered direction $\boldsymbol{\xi}$ in terms of the incident and scattered polar and azimuthal angles gives the scattering angle ψ:

$$\begin{aligned} \cos\psi &= \boldsymbol{\xi}' \cdot \boldsymbol{\xi} \\ &= \cos\theta' \cos\theta + \sin\theta' \sin\theta \cos(\phi - \phi') \,. \end{aligned} \tag{1.4}$$

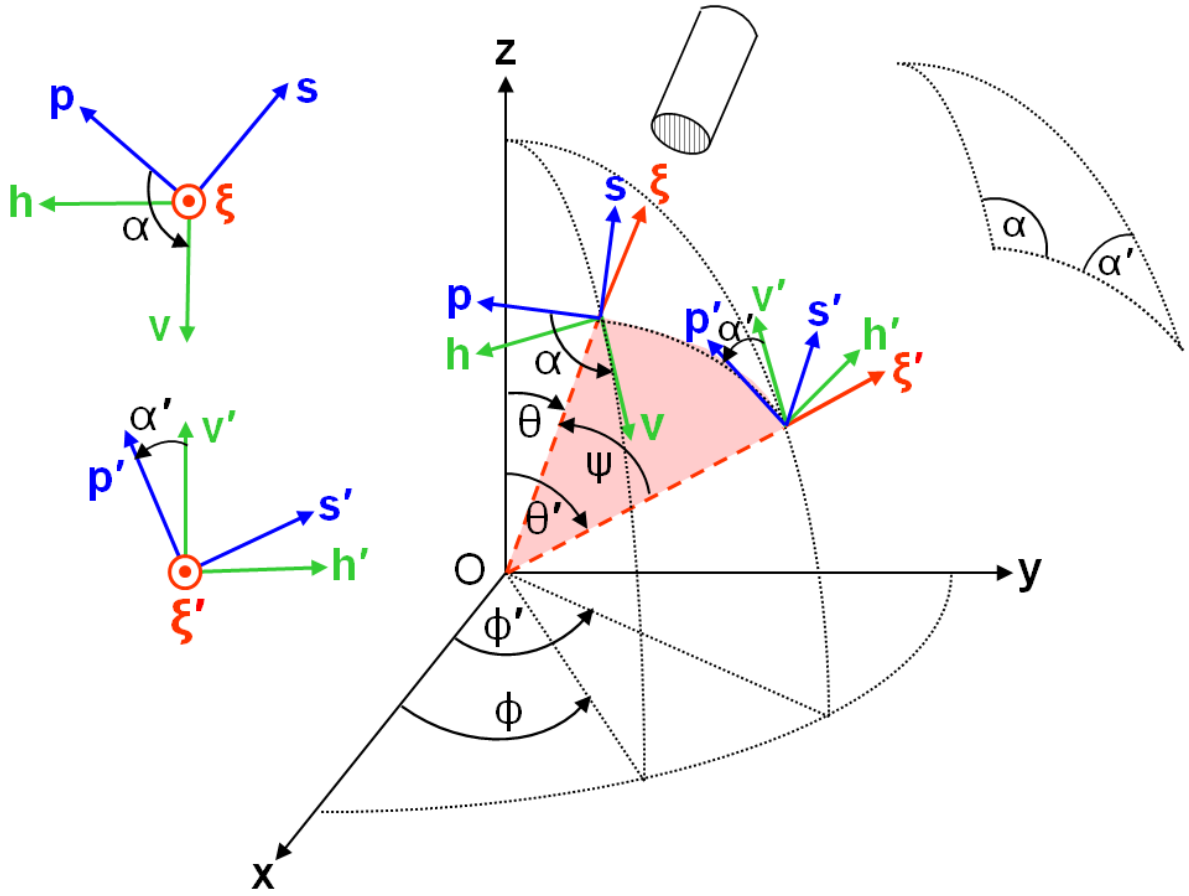

Figure 1.3: The scattering plane and the incident and final meridian planes showing the coordinate systems and rotations used to specify scattering of polarized light. The scattering plane is partly shaded in red.

Consider an incident beam of light propagating in direction $\boldsymbol{\xi}' = \mathbf{r}$ as in Fig. 1.1. Direction $\boldsymbol{\xi}'$ is specified by polar and azimuthal directions (θ', ϕ'). Primed variables denote unscattered or incident directions; unprimed variables denote scattered or final directions. The $\mathbf{z}$ axis and the direction of light propagation $\boldsymbol{\xi}'$ define the *incident meridian plane*, part of which is shaded in blue in the right panel of Fig. 1.1. The 4×1 (diffuse) Stokes vector $\underline{S}' = [I', Q', U', V']^T$ for this beam of light is

described with reference to "horizontal" and "vertical" directions, $+\mathbf{h}'$ and $+\mathbf{v}'$ respectively, which were defined above; the superscript T denotes transpose. Note that the horizontal unit vector $+\mathbf{h}'$ is perpendicular to the meridian plane, and the vertical vector $+\mathbf{v}'$ is parallel to the meridian plane.

To compute how an incident Stokes vector $\underline{S}'$ is scattered to a final vector $\underline{S}$, the horizontal and vertical components of $\underline{S}'$ in the incident meridian plane must first be rotated into components parallel and perpendicular to the scattering plane. The coordinate system after rotation of $\mathbf{v}'$ and $\mathbf{h}'$ about the $\boldsymbol{\xi}'$ axis is labeled $\mathbf{p}'$ (parallel to the scattering plane) and $\mathbf{s}'$ (perpendicular to the scattering plane). Note that $\mathbf{s}' \times \mathbf{p}'$ still gives the direction of propagation $\boldsymbol{\xi}'$. As shown in Fig. 1.3, rotation angle α' takes $\mathbf{v}'$ into $\mathbf{p}'$ (and $\mathbf{h}'$ into $\mathbf{s}'$). In HydroPol, *rotation angles are positive for counterclockwise rotations when looking "into the beam,"* e.g. in the $-\boldsymbol{\xi}'$ direction. This is similar to rotations about the $\mathbf{z}$ axis of Fig. 1.1 having positive angles ϕ for counterclockwise rotations when looking in the $-\mathbf{z}$ direction.

When computing single scattering with both $\boldsymbol{\xi}'$ and $\boldsymbol{\xi}$ being expressed in their respective meridian planes, the rotation angles can be obtained from spherical trigonometry applied to the triangle defined by $\mathbf{z}$, $\boldsymbol{\xi}'$, and $\boldsymbol{\xi}$, which is shown in the inset in Fig. 1.3. Given $\theta', \phi', \theta, \phi$, spherical trigonometry gives the rotation angles α' and α as (e.g., van de Hulst (1980), Vol. 2, page 499, or Mischenko et al. (2002) page 90)

$$\cos\alpha' = (\cos\theta - \cos\theta'\cos\psi)/(\sin\psi\sin\theta') \tag{1.5}$$

or

$$\sin\alpha' = -\sin\theta\sin(\phi - \phi')/\sin\psi\,. \tag{1.6}$$

and

$$\cos\alpha = (\cos\theta' - \cos\theta\cos\psi)/(\sin\psi\sin\theta) \tag{1.7}$$

or

$$\sin\alpha = -\sin\theta'\sin(\phi - \phi')/\sin\psi \tag{1.8}$$

for $\psi \neq 0$ or π and for $0 \leq \phi - \phi' \leq \pi$. If $\pi < \phi - \phi' < 2\pi$, then α' and α are given by the negatives of these equations. The scattering angle ψ is given by Eq. (1.4). Note that the rotation and scattering angles depend only on the difference in azimuthal angles via $\phi - \phi'$. Special cases are required when θ', θ, or ψ are zero. For $\psi = 0$, set $\alpha' = \alpha = 0$ since $\phi' = \phi \implies \sin(\phi - \phi') = 0$. For $\psi = \pi$, set $\alpha' = \alpha = 0$ since $\phi = \phi' + \pi \implies \sin(\phi - \phi') = 0$. If $\sin\theta' = 0$, replace Eqs. (1.5) and (1.7) with (Hu et al., 2001)

$$\cos\alpha' = -\cos\theta'\cos(\phi - \phi') \tag{1.9}$$

$$\cos\alpha = \cos\theta'\,. \tag{1.10}$$

If $\sin\theta = 0$, replace Eqs. (1.5) and (1.7) with

$$\cos\alpha' = \cos\theta \tag{1.11}$$

$$\cos\alpha = -\cos\theta\cos(\phi - \phi')\,. \tag{1.12}$$

However, when doing calculations of multiple scattering between sea surface wave facets, a light ray can reflect from one wave facet to another several times before the incident ray finally leaves the surface region and needs to be rotated into the final meridian plane. In this case, it is more

convenient to obtain the rotation angles from the perpendicular (or parallel) axes as determined for the incident ray direction onto a facet and the normal to the tilted wave facet. The details of these calculations are given in §3.2. However, it can be seen from Fig. 1.3 that the rotation angles can be obtained from $\cos^{-1}(\mathbf{v}' \cdot \mathbf{p}')$ and $\cos^{-1}(\mathbf{p} \cdot \mathbf{v})$, where the dot denotes the vector dot product. The exact equations that yield rotation angles from these vector dot products are given in §3.2.2.

Once the incident Stokes vector is specified in the scattering plane, the scattering matrix is applied to obtain the final Stokes vector, which is then expressed in the **s**-**p**-$\boldsymbol{\xi}$ scattering plane coordinate system defined for the final direction: $\mathbf{s} \times \mathbf{p} = \boldsymbol{\xi}$. Finally, the parallel and perpendicular components of the final Stokes vector must be expressed as horizontal and vertical components in the final meridian plane as specified by the **h**-**v**-$\boldsymbol{\xi}$ system. As illustrated in Fig. 1.3, this requires a counterclockwise rotation through an angle of α, where α is the "interior" angle of the spherical triangle illustrated in the figure. If we let $\underline{R}(\gamma)$ represent a counterclockwise (positive) rotation through angle γ and $\underline{M}(\psi)$ represent scattering through scattering angle ψ, then we can symbolically represent this scattering process by

$$\underline{S} = \underline{R}(\alpha)\underline{M}(\psi)\underline{R}(\alpha')\underline{S}' \,. \tag{1.13}$$

There is a subtlety in these rotations. Stokes vectors are not vectors in 3D space like the various unit vectors specifying directions, they are just 4-tuples of real numbers. (True vectors are defined by how their components behave under a coordinate transformation, and Stokes vectors don't satisfy these transformation conditions.) Two of the Stokes components, I and V are independent of the coordinate system. But the other two components, Q and U, are defined relative to the *plane* of vibration (not a unique direction) of the electric field vector and thus depend on the coordinate system used to specify the "horizontal" and "vertical" planes. The consequence of this is that we do not need to resolve the Stokes vectors along one particular direction, but only with reference to planes. This can be done with the **s** and **p** vectors as defined above, but using -**s** and -**p** would work equally well. Thus the counterclockwise rotation through angle α leading to $\mathbf{h} = -\boldsymbol{\varphi}$ and $\mathbf{v} = \boldsymbol{\vartheta}$ as seen in Fig. 1.3 is equivalent to a clockwise rotation by $\pi - \alpha$, leading to $\mathbf{h} = \boldsymbol{\varphi}$ and $\mathbf{v} = -\boldsymbol{\vartheta}$, as seen for $\boldsymbol{\xi}'$. This ambiguity means that the rotation matrices that carry Stokes vectors from one coordinate system to another depend on twice the rotation angle that carries true direction vectors from one coordinate system to another. That is to say, for a rotation of coordinate systems through an angle γ, the rotation matrix $\underline{R}(\gamma)$ depends on 2γ.

For our choice of a positive rotation being counterclockwise when looking into the beam, the Stokes vector rotation matrix is (e.g., Mischenko et al. (2002) page 25)

$$\underline{R}(\gamma) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos 2\gamma & -\sin 2\gamma & 0 \\ 0 & \sin 2\gamma & \cos 2\gamma & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \,. \tag{1.14}$$

These rotation matrices have several obvious but important properties:

$$\underline{R}(\pi) = \underline{I}\,, \tag{1.15a}$$
$$\underline{R}(\pi+\gamma) = \underline{R}(\gamma)\,, \tag{1.15b}$$
$$\underline{R}(\pi-\gamma) = \underline{R}(-\gamma)\,, \tag{1.15c}$$
$$\underline{R}(-\gamma) = \underline{R}^{-1}(\gamma) = \underline{R}^{T}(\gamma)\,, \tag{1.15d}$$
$$\underline{R}(\gamma_1)\underline{R}(\gamma_2) = \underline{R}(\gamma_1+\gamma_2)\,. \tag{1.15e}$$

Equation (1.15a) shows that rotating a coordinate system through an angle of π, which turns **s** and **p** into -**s** and -**p** leaves the Stokes vector unchanged, consistent with the previous remark about directions -**s**, -**p** being equivalent to **s**, **p** as regards Stokes vectors.

It is noted for completeness that a positive rotation in 3D space by angle γ about an axis specified by unit vector $\mathbf{u} = u_x\mathbf{x} + u_y\mathbf{y} + u_z\mathbf{z} = (u_x, u_y, u_x)$ is

$$\underline{R}_{3D}(\mathbf{u},\gamma) = \begin{bmatrix} \cos\gamma + u_x^2(1-\cos\gamma) & u_xu_y(1-\cos\gamma) - u_z\sin\gamma & u_xu_z(1-\cos\gamma) + u_y\sin\gamma \\ u_yu_x(1-\cos\gamma) + u_z\sin\gamma & \cos\gamma + u_y^2(1-\cos\gamma) & u_yu_z(1-\cos\gamma) - u_x\sin\gamma \\ u_zu_x(1-\cos\gamma) - u_y\sin\gamma & u_zu_y(1-\cos\gamma) + u_x\sin\gamma & \cos\gamma + u_z^2(1-\cos\gamma) \end{bmatrix} \tag{1.16}$$

for a counterclockwise rotation when looking in the $-\mathbf{u}$ direction. Although not needed for Stokes vector manipulations, this matrix can be used to check quantities determined by other means. For example, $\mathbf{s}' = \underline{R}_{3D}(\boldsymbol{\xi}', \alpha')\mathbf{h}'$ as seen in Fig. 1.3.

The choice of coordinate systems and rotation angles is not unique. Recall the definition of the Stokes vector in Eq. (1.1), which involves the components of the electric field in directions that are parallel and perpendicular to a "conveniently chosen" reference plane. Kattawar and Adams (1989), Kattawar (1994), Zhai et al. (2012), and Mischenko et al. (2002) all choose that reference plane to the be meridian plane. (These authors use somewhat different notation; our **h** and **v** are Kattawar's **r** and **l**, respectively. Mischenko et al. (2002) (page 16) on the other hand use the spherical coordinate system unit vectors $\boldsymbol{\vartheta}$ and $\boldsymbol{\varphi}$ for the vertical and horizontal axes.) Thus in the ocean setting they regard the "horizontal" direction (parallel to the mean sea surface) as being the "perpendicular" direction (relative to the meridian plane), and "vertical" to the mean sea surface as being the "parallel" direction. Note in Eq. (1.1) that if the parallel component of the electric field is zero, then the second element of the Stokes vector is proportional to $-E_\perp E_\perp^* = -E_\perp^2$, which is a negative number. A Stokes vector for light reflected from the sea surface with horizontal linear polarization is therefore proportional to $\underline{S} = [1, -1, 0, 0]^T$, which they refer to as perpendicular polarization. Their choice of the reference plane means that the Mueller matrix that transmits only horizontally polarized light has the form

$$\underline{\tilde{M}}_{\mathrm{h}} = \begin{bmatrix} 1 & -1 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}. \tag{1.17}$$

However, Bohren and Huffman (1983) and Hecht (1989) choose their "parallel" direction to be parallel to a horizontal direction, such as a laboratory bench top or the mean sea surface, and their

vertical direction is perpendicular to the bench top or mean sea surface. Thus their Stokes vector for horizontal polarization (parallel to the mean sea surface) is $\underline{S} = [1, 1, 0, 0]^T$ and their Mueller matrix that transmits only horizontally polarized light has the form

$$\underline{\tilde{M}}_{\mathrm{h}} = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} .$$

The difference choices arise perhaps from the viewpoints of describing polarization in a convenient way for a laboratory experiment with reference to a table top, versus modeling light incident onto the sea surface with reference to meridian planes.

Similar confusion is found in the choice of rotation angles. Kattawar (and his students in their papers) and Bohren and Huffman (1983) define a positive rotation as being *clockwise* when looking into the beam. Since a clockwise rotation through angle γ is the same as a counterclockwise rotation through $-\gamma$, Kattawar's rotation matrix is the transpose of the one in Eq. (1.14) (see Eq. 1.15d). Thus Kattawar (e.g., in Kattawar and Adams (1989) Eq. (10)) writes Eq. (1.13) as $\underline{S} = \underline{R}(-\alpha)\underline{M}(\psi)\underline{R}(-\alpha')\underline{S}'$ (again, with minor differences in notation; my α' is Kattawar's Φ, etc.). Others often write a rotation as $R(\pi - \alpha)$ rather than $R(-\alpha)$; these are equivalent by Eq. (1.15c). Chandrasekhar (1960) also defines a positive rotation as being clockwise when looking into the beam. However, he uses a different definition for the Stokes vector for which only the fourth component is independent of coordinate system, so his rotation matrix is more complicated. Mischenko defines a positive rotation as being clockwise when looking in the direction of propagation. This is equivalent to counterclockwise when looking into the beam as used here; thus his rotation matrix is the same as that in Eq. (1.14). van de Hulst (1980) also uses the same rotation convention as is used in HydroPol. All of this rotation business is considered well known, so the details are often omitted in publications. ("Well known," of course, means that when you encounter this material for the first time, you have to spend many days figuring out the confusing apparent differences in publications by different authors.) Fortunately, the only real requirement for Stokes vectors, coordinate systems, and rotations is consistency in usage once a choice has been made.

1.3 The Vector Radiative Transfer Equation

With this understanding of how to describe scattering of polarized light, we can upgrade the SRTE to the vector level. Both HydroLight and HydroPol use a plane-parallel ocean in which the inherent optical properties (IOPs) can vary arbitrarily with depth but do not vary with horizontal position. The same holds true of the surface and bottom boundary conditions; they are the same for all horizontal positions. These codes therefore cannot simulate sloping bottoms, the effects of objects in the water column, point sources of bioluminescence, or the underwater light from an airborne laser with a finite surface footprint; these are all inherently 3D radiative transfer problems for which Monte Carlo simulation is the appropriate solution technique. Likewise, HydroLight and HydroPol solve time-independent RTEs and cannot simulate time-dependent processes such as pulse stretching from a pulsed lidar.

The time-independent, 1D, SRTE is (L&W 5.23)

$$\begin{aligned}\cos\theta\frac{dL(z,\theta,\phi,\lambda)}{dz} &= -c(z,\lambda)L(z,\theta,\phi,\lambda)\\ &\quad+\int_0^{2\pi}\int_0^{\pi}\beta(z;\theta',\phi'\to\theta,\phi;\lambda)L(z,\theta',\phi',\lambda)\sin\theta' d\theta' d\phi'\\ &\quad+\sigma(z,\theta,\phi,\lambda)\,. \end{aligned}\tag{1.18}$$

Here L is the unpolarized radiance and σ is a source term that represents radiance inputs from bioluminescence, Raman scatter, or fluorescence; σ has units of $\mathrm{W\,m^{-3}\,sr^{-1}\,nm^{-1}}$. This equation expresses location as geometric depth z in meters and the IOPs in terms of the beam attenuation c and the volume scattering function β. Note that β has units of $\mathrm{m^{-1}\,sr^{-1}}$. The notation $\beta(z;\theta',\phi'\to\theta,\phi;\lambda)$ reminds us that we are scattering radiance from incident direction (θ',ϕ') to final direction (θ,ϕ). The factor $\sin\theta' d\theta' d\phi'$ is the differential element of solid angle in spherical coordinates.

In this work, we consider only *isotropic media*, which means that the inherent optical properties are independent of direction within the medium[1]. This in turn means that scattering depends only on the angle between the incident and scattered directions, not on the directions themselves, as seen in Eq. (1.4).

The nondimensional *optical depth* ζ is defined by

$$d\zeta(\lambda) = c(z,\lambda)dz\,.\tag{1.19}$$

This equation can be integrated to convert geometric depth to optical depth, or vice versa:

$$\zeta(z,\lambda) = \int_0^z c(z',\lambda)dz' \quad \text{or} \quad z(\zeta,\lambda) = \int_0^{\zeta}\frac{d\zeta'}{c(\zeta',\lambda)}\,.\tag{1.20}$$

Note that the optical depth corresponding to a given physical depth z depends on wavelength via $c(z,\lambda)$.

Dividing Eq. (1.18) by $c(z,\lambda)$ and using Eq. (1.19) gives the SRTE in terms of optical depth. It is common to use $\mu=\cos\theta$ as the polar angle variable. We can factor the volume scattering function β into the scattering coefficient b (units of $\mathrm{m^{-1}}$) times the scattering phase function $\tilde{\beta}$ (units of $\mathrm{sr^{-1}}$). Finally, defining the non-dimensional *albedo of single scattering* $\omega_\mathrm{o}=b/c$, we can re-write Eq. (1.18) as

$$\begin{aligned}\mu\frac{dL(\zeta,\mu,\phi,\lambda)}{d\zeta} &= -L(\zeta,\mu,\phi,\lambda)\\ &\quad+\omega_\mathrm{o}(\zeta,\lambda)\int_0^{2\pi}\int_{-1}^{1}\tilde{\beta}(\zeta;\mu',\phi'\to\mu,\phi;\lambda)L(\zeta,\mu',\phi',\lambda)d\mu' d\phi'\\ &\quad+\frac{1}{c(\zeta,\lambda)}\sigma(\zeta,\mu,\phi,\lambda)\,. \end{aligned}\tag{1.21}$$

All quantities are now shown as functions of optical depth.

[1]Crystals, or a cirrus cloud containing oriented ice crystals, are examples of anisotropic media that have different optical properties for different directions within the medium.

To upgrade Eq. (1.18) to the vector level, we need to

1. Replace the radiance L with the Stokes vector $\underline{S}$;
2. Replace the volume scattering function β with the *phase matrix* $\underline{P}$;
3. Replace the source term for scalar radiance, σ, with one for the emission of polarized light.

Note that the 4×4 phase matrix $\underline{P}$, which is the vector equivalent of the volume scattering function, also has units of $\mathrm{m}^{-1}\,\mathrm{sr}^{-1}$.

To upgrade Eq. (1.21) to the vector level, we need to

1. Replace the radiance L with the Stokes vector $\underline{S}$;
2. Replace the phase function $\tilde{\beta}$ with the *normalized phase matrix* $\underline{\tilde{P}} = \underline{P}/b$;
3. Replace the source term for scalar radiance, σ, with one for the emission of polarized light, $\underline{\Sigma}$.

Both the phase function $\tilde{\beta}$ and the reduced phase matrix $\underline{\tilde{P}}$ have units of sr^{-1}. The 1D, time-independent, VRTE is thus compactly written in matrix form as

$$\begin{aligned}\mu\frac{d\underline{S}(\zeta,\mu,\phi,\lambda)}{d\zeta} &= -\,\underline{S}(\zeta,\mu,\phi,\lambda)\\ &\quad+\,\omega_{\mathrm{o}}(\zeta,\lambda)\int_0^{2\pi}\int_0^{\pi}\underline{\tilde{P}}(\zeta;\mu',\phi'\rightarrow\mu,\phi;\lambda)\,\underline{S}(\zeta,\mu',\phi',\lambda)'d\mu'd\phi'\\ &\quad+\,\underline{\Sigma}(\zeta,\mu,\phi,\lambda)\,. \end{aligned}\tag{1.22}$$

This matrix equation represents 4 scalar equations for the 4 components of the Stokes vector. The integral term connects these four components and shows how scattering transforms various states of polarization (components of the Stokes vector) for light traveling in all incident directions into other states of polarization of light traveling in the θ,ϕ direction of interest. In particular, the (i,j) element of $\underline{\tilde{P}}$ shows how scattering converts incident light in polarization state j (the j^{th} element of $\underline{S}'$) into polarization state i of the final Stokes vector $\underline{S}$. The $1/c(\zeta,\lambda)$ factor multiplying the source function σ in the scalar equation (1.21) has for notational convenience been incorporated into the definition of the vector source function $\underline{\Sigma}$, which therefore has units of radiance, $\mathrm{W\,m^{-2}\,sr^{-1}\,nm^{-1}}$.

1.4 The Phase Matrix

We next consider the phase matrix in more detail. The phase matrix can be expressed in various ways:

$$\begin{aligned}\underline{P}(\zeta;\mu',\phi',\rightarrow\mu,\phi;\lambda) &= \underline{P}(\zeta;\mu',\mu,\phi-\phi';\lambda)\\ &= \underline{R}(\alpha)\,\underline{M}(\psi)\,\underline{R}(\alpha') &(1.23)\\ &= \beta(\psi)\,\underline{R}(\alpha)\,\underline{\tilde{M}}(\psi)\,\underline{R}(\alpha') &(1.24)\\ &= b(\zeta,\lambda)\,\tilde{\beta}(\psi)\,\underline{R}(\alpha)\,\underline{\tilde{M}}(\psi)\,\underline{R}(\alpha') &(1.25)\\ &= b(\zeta,\lambda)\,\tilde{\beta}(\psi)\,\underline{\tilde{\tilde{P}}} &(1.26)\\ &= b(\zeta,\lambda)\,\underline{\tilde{P}}\,. &(1.27)\end{aligned}$$

The terminology relating to phase matrices is confusing, and there is no uniformity of notation in the literature. It is therefore worth spelling out several distinctions:

- The *phase matrix* $\underline{P}$ is a 4×4 matrix that transforms the incident or unscattered Stokes vector $\underline{S}'$ into the final or scattered vector $\underline{S}$, *with both vectors expressed in meridian planes.*
- The scattering matrix $\underline{M}$ is a 4×4 matrix that transforms $\underline{S}'$ into $\underline{S}$, *with both vectors expressed in the scattering plane.* (The term scattering matrix is commonly used when the scattering is caused by particles or molecules, as in the ocean. In laboratory optics the transformation from $\underline{S}'$ to $\underline{S}$ is often caused by lenses, filters, mirrors, etc, and the term Mueller matrix is commonly used. However, scattering matrix and Mueller matrix are synonymous in that they both transform Stokes vectors in the scattering plane.) The scattering matrix elements generally depend on depth and wavelength, not shown here, as well as on scattering angle.
- $\underline{R}(\alpha')$ and $\underline{R}(\alpha)$ are the *rotation matrices* that carry Stokes vectors into and out of the scattering plane. The rotation matrices are non-dimensional; it is the scattering matrix that carries the units of $\mathrm{m}^{-1}\,\mathrm{sr}^{-1}$, and the phase matrix therefore has the same units.
- The (1,1) element of the scattering matrix, $M_{1,1}$, transforms the total incident radiance, the I' element of $\underline{S}'$, into total scattered radiance, the I element of $\underline{S}$. In different terminology, $M_{1,1}(\psi)$ is thus the volume scattering function $\beta(\psi)$ as used in scalar radiative transfer theory.
- It is customary to factor the volume scattering function $M_{1,1}(\psi)$ out of each element of $\underline{M}$ to obtain a dimensionless *reduced scattering matrix* $\underline{\tilde{M}}$ whose elements are $\tilde{M}_{i,j}(\psi) = M_{i,j}(\psi)/M_{1,1}(\psi)$. This is similar to factoring the scattering coefficient $b(z,\lambda)$ out of the volume scattering function to obtain the *scattering phase function* $\tilde{\beta}(\psi) = \beta(\psi)/b(z,\lambda)$, as seen in Eq. (1.25).
- $\underline{\tilde{\tilde{P}}}$ is called the *reduced phase matrix* because it is constructed from the reduced scattering matrix $\underline{\tilde{M}}$; it is likewise dimensionless.
- $\underline{\tilde{P}}$ in the last equation is called the *normalized phase matrix.* Note that the (1,1) element of the normalized phase matrix $\underline{\tilde{P}}$ is just the scattering phase function $\tilde{\beta}$. $\underline{\tilde{P}}$ therefore has units of sr^{-1}.

By definition, the phase function $\tilde{\beta}(\theta', \phi' \to \theta, \phi)$ gives the probability per unit solid angle Ω centered on (θ, ϕ) that radiance traveling in any direction (μ', ϕ') will be scattered into any other direction (μ, ϕ). The phase function therefore satisfies the normalization condition

$$\int_{\phi=0}^{2\pi} \int_{\theta=0}^{\pi} \tilde{\beta}(\theta', \phi' \to \theta, \phi)) d\Omega(\theta, \phi) = 1 \qquad \text{for any } (\theta', \phi')\,. \tag{1.28}$$

That is, the probability is 1 that the radiance will be scattered into some direction. For isotropic media, the phase function depends only on the scattering angle ψ and this equation can be written

$$2\pi \int_0^{\pi} \tilde{\beta}(\psi) \sin\psi d\psi = 2\pi \int_0^{\pi} \tilde{P}_{1,1}(\psi) \sin\psi d\psi = 1\,. \tag{1.29}$$

For almost all problems in optical oceanography, it is sufficient to consider scattering by a dilute mixture of randomly oriented, non-spherical particles such as phytoplankton or minerals. (Here

"dilute" means that the individual particles are separated by many times the wavelength of the light, so that coherent scattering does not occur.) Symmetry considerations for such particles (e.g. Mischenko et al., 2002) show that the scattering matrix then has only 10 independent components. The reduced scattering matrix for such particles has the form

$$\underline{\tilde{M}} = \begin{bmatrix} 1 & \tilde{M}_{12}(\psi) & \tilde{M}_{13}(\psi) & \tilde{M}_{14}(\psi) \\ \tilde{M}_{12}(\psi) & \tilde{M}_{22}(\psi) & \tilde{M}_{23}(\psi) & \tilde{M}_{24}(\psi) \\ -\tilde{M}_{13}(\psi) & -\tilde{M}_{23}(\psi) & \tilde{M}_{33}(\psi) & \tilde{M}_{34}(\psi) \\ \tilde{M}_{14}(\psi) & \tilde{M}_{24}(\psi) & -\tilde{M}_{34}(\psi) & \tilde{M}_{44}(\psi) \end{bmatrix} . \tag{1.30}$$

The 10^{th} independent element of the scattering matrix $\underline{M}$ is the the volume scattering function $\underline{M}_{1,1}(\psi) = \beta(\psi)$, which has been factored out in creating $\underline{\tilde{M}}$. Particles with additional symmetries, such as mirror-symmetric particles or spherical particles, have simpler scattering matrices with many of the above matrix elements being zero. However, HydroPol allows for matrices of the form of Eq. (1.30), even though the $\underline{\tilde{M}}$ used in a particular calculation is usually simpler.

The normalized phase matrix $\underline{\tilde{P}}$ as used in HydroPol (Eq. 1.22) can now be written in full as

$$\begin{aligned} \underline{\tilde{P}} &= \tilde{\beta}\, \underline{R}(\alpha)\, \underline{\tilde{M}}(\psi)\, \underline{R}(\alpha') \\ &= \tilde{\beta}(\psi) \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos 2\alpha & -\sin 2\alpha & 0 \\ 0 & \sin 2\alpha & \cos 2\alpha & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \\ &\times \begin{bmatrix} 1 & \tilde{M}_{12}(\psi) & \tilde{M}_{13}(\psi) & \tilde{M}_{14}(\psi) \\ \tilde{M}_{12}(\psi) & \tilde{M}_{22}(\psi) & \tilde{M}_{23}(\psi) & \tilde{M}_{24}(\psi) \\ -\tilde{M}_{13}(\psi) & -\tilde{M}_{23}(\psi) & \tilde{M}_{33}(\psi) & \tilde{M}_{34}(\psi) \\ \tilde{M}_{14}(\psi) & \tilde{M}_{24}(\psi) & -\tilde{M}_{34}(\psi) & \tilde{M}_{44}(\psi) \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos 2\alpha' & -\sin 2\alpha' & 0 \\ 0 & \sin 2\alpha' & \cos 2\alpha' & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \end{aligned} \tag{1.31}$$

Letting $c = \cos 2\alpha$, $s' = \sin 2\alpha'$, etc., and $m_{ij} = \tilde{M}_{i,j}(\psi)$, the normalized phase matrix after matrix multiplications becomes

$$\underline{\tilde{P}} = \tilde{\beta}(\psi) \begin{bmatrix} 1 & c'\, m_{12} + s'\, m_{13} & -s'\, m_{12} + c'\, m_{13} & m_{14} \\ c\, m_{12} + s\, m_{13} & c'(c\, m_{22} + s\, m_{23}) + s'(c\, m_{23} - s\, m_{33}) & -s'(c\, m_{22} + s\, m_{23}) + c'(c\, m_{23} - s\, m_{33}) & c\, m_{24} - s\, m_{34} \\ s\, m_{12} - c\, m_{13} & c'(s\, m_{22} - c\, m_{23}) + s'(s\, m_{23} + c\, m_{33}) & -s'(s\, m_{22} - c\, m_{23}) + c'(s\, m_{23} + c\, m_{33}) & s\, m_{24} + c\, m_{34} \\ m_{14} & c'\, m_{24} - s'\, m_{34} & -s'\, m_{24} - c'\, m_{34} & m_{44} \end{bmatrix} \tag{1.32}$$

Note that although the scattering matrix elements $\tilde{M}_{ij}(\psi)$ and $\tilde{M}_{ji}(\psi)$ have the symmetries seen in Eq. (1.30), those symmetries are lost in the phase matrix because of the rotations.

1.4.1 The Rayleigh Scattering Matrix

If the scattering particles are homogeneous spheres that are much smaller than the wavelength of light, the so-called Rayleigh scattering regime, the scattering matrix has the form

$$\begin{aligned}\underline{M}_{\text{Ray}} =&\beta_{\text{Ray}}\underline{\tilde{M}}_{\text{Ray}}\\ =&b_{\text{Ray}}\tilde{\beta}_{\text{Ray}}\underline{\tilde{M}}_{\text{Ray}}\\ =&b_{\text{Ray}}\frac{3}{16\pi}(1+\cos^2\psi)\begin{bmatrix}1 & -\frac{\sin^2\psi}{1+\cos^2\psi} & 0 & 0\\ -\frac{\sin^2\psi}{1+\cos^2\psi} & 1 & 0 & 0\\ 0 & 0 & \frac{2\cos\psi}{1+\cos^2\psi} & 0\\ 0 & 0 & 0 & \frac{2\cos\psi}{1+\cos^2\psi}\end{bmatrix}\end{aligned} \tag{1.33}$$

Note that the Rayleigh phase function $\tilde{\beta}_{\text{Ray}} = \frac{3}{16\pi}(1+\cos^2\psi)$ satisfies the normalization condition (1.29). Figure 1.4 shows the reduced Rayleigh scattering matrix plotted as a 4×4 plot, with the upper left plot showing the (1,1) element of $\underline{\tilde{M}}_{\text{Ray}}$ and the lower right plot showing $\tilde{M}_{\text{Ray}}(4,4)$.

Measurements of scattering matrices for ocean water (Voss and Fry, 1984) and different algal species (Svensen et al., 2011) show that the reduced scattering matrices are similar to that for Rayleigh scattering, although the phase functions are much more highly peaked at small scattering angles than $\tilde{\beta}_{\text{Ray}}$. Thus, to a first approximation, scattering matrices for ocean water can be modeled by combining a highly peaked phase function with a Rayleigh reduced scattering matrix:

$$\begin{aligned}\underline{M} =&\beta\,\underline{\tilde{M}}_{\text{Ray}}\\ =&b\,\tilde{\beta}\,\underline{\tilde{M}}_{\text{Ray}}\,.\end{aligned} \tag{1.34}$$

Such scattering matrices are sufficient for code debugging and idealized simulations. The reduced Rayleigh scattering matrix can be replaced by measured values for more realistic simulations.

1.4.2 Dependence of the Phase Matrix on Azimuthal Angle

The phase function $\tilde{\beta}(\psi)$ in the SRTE depends on the scattering angle ψ, computed via Eq. (1.4). This equation shows that the phase function depends on the incident and scattered azimuthal angles via cosine of the difference in the azimuthal angles. Because $\tilde{\beta}(\mu',\phi'\rightarrow\mu,\phi)$ depends only on the difference $\phi-\phi'$, we can without loss of generality set $\phi'=0$, which merely anchors the difference $\phi-\phi'$ to $\phi'=0$ for computational purposes. The dependence on $\cos(\phi-\phi')$ also means that the phase function can be expanded as a Fourier cosine series in the azimuthal angle ϕ.

Unfortunately this $\cos(\phi-\phi')$ dependence is lost in the vector phase matrix because of the rotation matrices (van de Hulst, 1980). The phase matrix still depends on $\phi-\phi'$, but some phase matrix elements depend on $\cos(\phi-\phi')$ and some depend on $\sin(\phi-\phi')$. Figure 1.5 illustrates this for the normalized Rayleigh phase matrix defined in Eqs. (1.27) and (1.33). Note that the non-zero elements of the upper left and lower right 2×2 blocks of matrix elements are non-zero at $\phi-\phi'=0$ and 180 deg. These 8 matrix elements can be expanded as Fourier cosine series, just as for the scalar phase function. However, the non-zero elements of the lower left and upper right 2×2 blocks of elements are 0 at $\phi-\phi'=0$ and 180 deg. These matrix elements therefore must be expanded as Fourier sine series. This result holds true for any scattering matrix. This mixed cosine and sine and

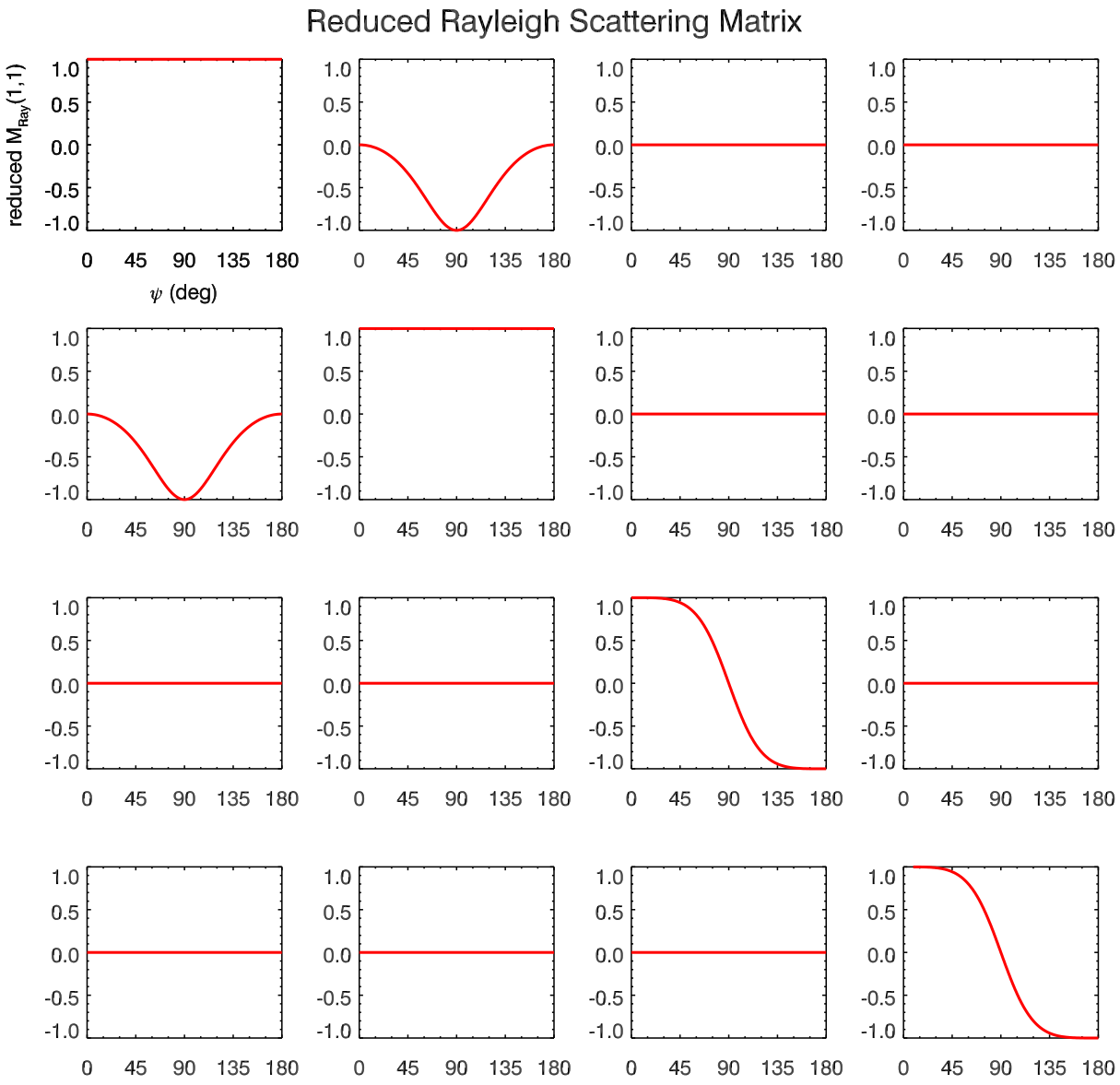

Figure 1.4: The reduced Rayleigh scattering matrix $\underline{\tilde{M}}_{\mathrm{Ray}}(\psi)$. Each element is plotted as a function of the scattering angle ψ. The upper left plot is the (1,1) matrix element; the lower right plot is the (4,4) element.

dependence of the phase matrix elements on azimuthal angle considerably complicates the solution of the VRTE compared to that of the SRTE, as will be seen in Chapter 4.

The phase function for the SRTE can be written as a function of either $\cos(\phi-\phi')$ or $\cos(\phi'-\phi)$. However, reversing the order of ϕ and ϕ' in the phase matrix of the VRTE introduces a sign change in the elements that depend on $\sin(\phi-\phi')$. For consistency, the development below always writes azimuthal angles in the order of scattered direction minus incident direction, i.e. $\phi-\phi'$, or $\phi_v-\phi_s$ after discretization.

1.4.3 Symmetries of the Phase Matrix

Hovenier (1969) derived several symmetry relations that must be obeyed by phase matrices. These symmetries result from considerations of time-reversal invariance of the scattering process and from the geometric symmetries of the scattering particles. Define the matrices

$$\underline{\Pi} = \underline{\Pi}^{-1} = \underline{\Pi}^T = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

and

$$\underline{\Upsilon} = \underline{\Upsilon}^{-1} = \underline{\Upsilon}^T = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix}.$$

Phase matrices whose scattering matrix describes randomly oriented non-spherical particles, i.e. whose normalized phase matrices have form of Eq. (1.32), obey the symmetry relations

$$\underline{P}(-\mu', -\mu, \phi'-\phi) = \underline{P}(\mu', \mu, \phi-\phi') \tag{1.35a}$$

$$\underline{P}(-\mu, -\mu', \phi'-\phi) = \underline{\Pi}\,\underline{P}^T(\mu', \mu, \phi-\phi')\,\underline{\Pi} \tag{1.35b}$$

$$\underline{P}(\mu', \mu, \phi'-\phi) = \underline{\Pi}\,\underline{\Upsilon}\,\underline{P}(\mu', \mu, \phi-\phi')\,\underline{\Upsilon}\,\underline{\Pi} \tag{1.35c}$$

If the scattering matrix describes particles with a plane of symmetry, the reduced scattering matrix has the form

$$\underline{\tilde{M}} = \begin{bmatrix} 1 & \tilde{M}_{12}(\psi) & 0 & 0 \\ \tilde{M}_{12}(\psi) & \tilde{M}_{22}(\psi) & 0 & 0 \\ 0 & 0 & \tilde{M}_{33}(\psi) & \tilde{M}_{34}(\psi) \\ 0 & 0 & -\tilde{M}_{34}(\psi) & \tilde{M}_{44}(\psi) \end{bmatrix}. \tag{1.36}$$

In that case, the phase matrix obeys four additional symmetry relations:

$$\underline{P}(\mu, \mu', \phi'-\phi) = \underline{\Upsilon}\,\underline{P}^T(\mu', \mu, \phi-\phi')\,\underline{\Upsilon} \tag{1.37a}$$

$$\underline{P}(\mu, \mu', \phi-\phi') = \underline{\Pi}\,\underline{P}^T(\mu', \mu, \phi-\phi')\,\underline{\Pi} \tag{1.37b}$$

$$\underline{P}(-\mu, -\mu', \phi-\phi') = \underline{\Upsilon}\;\underline{P}^T(\mu', \mu, \phi-\phi')\,\underline{\Upsilon} \tag{1.37c}$$

$$\underline{P}(-\mu', -\mu, \phi-\phi') = \underline{\Pi}\,\underline{\Upsilon}\;\underline{P}(\mu', \mu, \phi-\phi')\,\underline{\Upsilon}\,\underline{\Pi} \tag{1.37d}$$

These symmetry relations provide an important check on the correctness of the discretized phase matrices, which will be computed as described in §2.2.

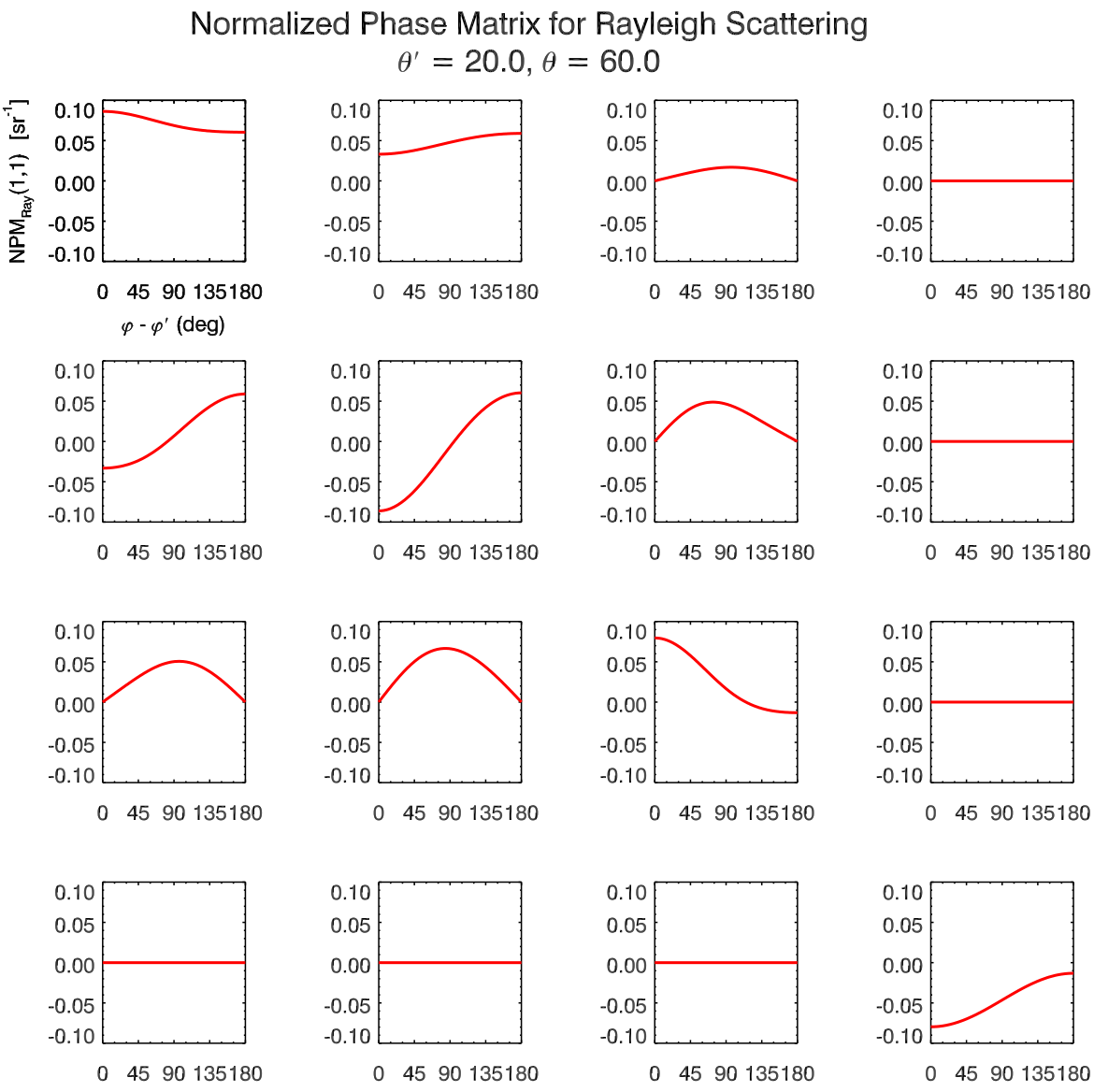

Figure 1.5: The normalized phase matrix for Rayleigh scattering, $\underline{\tilde{P}}_{\mathrm{Ray}}$, plotted as a function of $\phi - \phi'$ for fixed polar angles $\theta' = 20\,\mathrm{deg}$ and $\theta = 60\,\mathrm{deg}$. Note the block structure as regards expansion of elements in either cosine or sine series.

1.5 Boundary Conditions at the Sea Surface

The VRTE developed in the previous section describes how light is absorbed and scattered within the water. It is also necessary to determine how light enters and leaves the water, and how it is reflected by the bottom if the water has a finite depth. Thus, in order to solve the VRTE, we must specify boundary conditions for the Stokes vector at the sea-surface and for an opaque but reflecting surface at the sea bottom.

The boundary conditions at the sea surface for the scalar RTE are given by (L&W 4.3)

$$L(a,\theta,\phi) = \iint_{2\pi_d} r_{aw}(\theta',\phi' \to \theta,\phi)\, L(a,\theta',\phi')\, d\Omega(\theta',\phi') + \iint_{2\pi_u} t_{wa}(\theta',\phi' \to \theta,\phi)\, L(w,\theta',\phi')\, d\Omega(\theta',\phi') \quad \text{for } (\theta,\phi) \in 2\pi_{\mathrm{u}} \tag{1.38}$$

and (L&W 4.4)

$$L(w,\theta,\phi) = \iint_{2\pi_u} r_{wa}(\theta',\phi' \to \theta,\phi)\, L(w,\theta',\phi')\, d\Omega(\theta',\phi') + \iint_{2\pi_d} t_{aw}(\theta',\phi' \to \theta,\phi)\, L(a,\theta',\phi')\, d\Omega(\theta',\phi') \quad \text{for } (\theta,\phi) \in 2\pi_{\mathrm{d}} \tag{1.39}$$

The notation $\iint_{2\pi_d}$ ($\iint_{2\pi_u}$) denotes integration over all θ', ϕ' in the downward (upward) hemisphere of directions (2π steradians). The four *scalar radiance transfer functions* r_{aw}, r_{wa}, t_{aw} and t_{wa} have units of sr^{-1}. $r_{aw}(\theta',\phi' \to \theta,\phi)$, for example, specifies how radiance incident onto the sea surface from downwelling direction (θ',ϕ') is reflected back upward into direction (θ,ϕ) by the water surface, per unit of solid angle. Equation (1.38) thus shows that the upwelling radiance in the air (depth argument a) just above the sea surface comes from downwelling radiance in the air reflected back upward by the sea surface, and from upwelling radiance incident onto the sea surface from below (depth w) and transmitted through the surface. A similar interpretation holds for Eq. (1.39).

These equations are easily upgraded to the vector level. Just as for scattering within the water, the "surface scattering" processes of reflection and refraction are described by 4×4 scattering matrices. In parallel with the scalar case, there are four of these matrix *sea-surface transfer functions*: $\underline{r}_{aw}$ describes how air-incident light is reflected by the surface back to the air, $\underline{t}_{aw}$ describes how air-incident light is transmitted through the surface into the water, $\underline{r}_{wa}$ reflects water-incident light back to the water, and $\underline{t}_{wa}$ transmits light from the water into the air. The vector equivalents of (1.38) and 1.39) are thus

$$\underline{S}(a,\theta,\phi) = \iint_{2\pi_d} \underline{r}_{aw}(\theta',\phi' \to \theta,\phi)\, \underline{S}(a,\theta',\phi')\, d\Omega(\theta',\phi') + \iint_{2\pi_u} \underline{t}_{wa}(\theta',\phi' \to \theta,\phi)\, \underline{S}(w,\theta',\phi')\, d\Omega(\theta',\phi') \quad \text{for } (\theta,\phi) \in 2\pi_{\mathrm{u}} \tag{1.40}$$

and

$$\underline{S}(w,\theta,\phi) = \iint_{2\pi_u} \underline{r}_{wa}(\theta',\phi' \to \theta,\phi)\, \underline{S}(w,\theta',\phi')\, d\Omega(\theta',\phi') + \iint_{2\pi_d} \underline{t}_{aw}(\theta',\phi' \to \theta,\phi)\, \underline{S}(a,\theta',\phi')\, d\Omega(\theta',\phi') \quad \text{for } (\theta,\phi) \in 2\pi_{\mathrm{d}} \tag{1.41}$$

When solving the VRTE, the sky radiance $\underline{S}(a, \theta', \phi')$ incident onto the sea surface is considered known.

The four sea-surface transfer functions are conceptually like the phase matrix seen in the VRTE of Eq. (1.22). That is, they describe reflection or transmission (i.e., scattering) from incident direction (θ', ϕ') to final direction (θ, ϕ), with incident and final Stokes vectors expressed in their meridian planes. For a level sea surface, the incident and final meridian planes are the same. For a wind-blown sea surface, the reflected and transmitted light for a single incident ray will generally lie in a different meridian plane than the incident light. Indeed, for a time- or space-averaged collection of random sea surfaces, a given incident direction yields reflected and transmitted light in all directions. Chapter 3 shows how to compute these sea-surface transfer functions, first for a level surface, and then for wind-blown surfaces.

1.6 Bottom Boundary Conditions

For complete generality, the bottom boundary condition could be formulated with a pair of equations like those at the sea surface. This would allow, for example, modeling a fish tank with a clear glass bottom with light entering the water column from below the physical bottom. However, for problems of oceanographic interest, it is sufficient to consider a bottom that reflects light but does not allow light to enter the water from below the bottom. In that case, there is only radiance incident onto the bottom being reflected back upward into the water column. There is no radiance being transmitted upward through the bottom, nor are we interested in radiance being transmitted downward below the bottom. Thus the bottom boundary condition reduces to just

$$\underline{S}(m, \theta, \phi) = \iint_{2\pi_d} \underline{r}_{mb}(\theta', \phi' \rightarrow \theta, \phi)\, \underline{S}(m, \theta', \phi')\, d\Omega(\theta', \phi') \quad \text{for } (\theta, \phi) \in 2\pi_{\mathrm{u}}\,. \tag{1.42}$$

where depth argument m indicates the depth where the bottom boundary condition is applied. Recall from Fig. 1.2 that this is the maximum depth to which the VRTE is to be solved. Depth m can be either a finite depth with a physical bottom, or a depth in the water column below which the water in infinitely deep, homogeneous, and source-free. $\underline{r}_{mb}$ is a 4×4 matrix that describes the reflectance properties of the bottom layer $[m, b]$.

It should be noted that it does not matter whether $\underline{r}_{mb}$ describes a finite-depth, opaque, physical bottom (i.e, a layer $[m, b] = [m, m]$) or an infinitely deep water column (i.e, a layer $[m, b] = [m, \infty]$), so long as $\underline{r}_{mb}$ describes how downwelling light at depth m is reflected back upward by whatever is at depth m and below.

1.6.1 Finite-depth Bottoms

In the case of an opaque bottom at depth $z = m = b$, the bottom boundary condition is commonly written in terms of the vector bidirectional reflectance distribution function (VBRDF):

$$\underline{S}(m, \theta, \phi) = \iint_{2\pi_d} \underline{VBRDF}(\theta', \phi' \rightarrow \theta, \phi)\, \cos\theta'\, \underline{S}(m, \theta', \phi')\, d\Omega(\theta', \phi') \tag{1.43}$$

Each (i, j) element of the VBRDF has the form of a scalar BRDF that reflects light from incident direction (θ', ϕ') into direction (θ, ϕ), and from polarization state j to i. The (1,1) element of

the VBRDF is the scalar BRDF. In general, these matrix elements will depend on wavelength. The $\cos\theta'$ factor appears in this equation because the radiance reflectance function $\underline{r}_{mb}$ as used in HydroPol is defined for surfaces normal to the directions of light propagation, whereas $\underline{VBRDF}$ is defined for light incident onto a horizontal surface. Just as for the sea surface, $\underline{r}_{mb}$ and $\underline{VBRDF}$ have units of sr^{-1} and include the effects of rotations between meridian planes and the reflection plane.

Svensen et al. (2012) measured and modeled the reflectance properties of Spectralon for polarized light. Some models have been developed for the VBRDF of terrestrial vegetation and soils (e.g. Schott, 1999). However, there are no such measurements or models for typical ocean bottom materials such as sand, mud, sea grass, or coral. Such materials are typically highly scattering rough surfaces that diffusely reflect light into all directions. For most oceanographic applications, it will be sufficient to consider VBRDFs that depend only on the difference in reflected and incident azimuthal angles, i.e.,

$$\underline{VBRDF}(\theta', \phi' \rightarrow \theta, \phi) = \underline{VBRDF}(\theta', \theta, \phi - \phi') \,. \tag{1.44}$$

Such VBRDFs should describe most bottom materials, and HydroPol is constructed to handle any VBRDF of this form.

However, such VBRDFs would not describe a sandy bottom with a linear ripple structure, which would reflect light differently for "along ripple" vs "cross ripple" directions. Such bottoms could be modeled in the same way as the sea surface reflectance.

Multiple scattering tends to depolarize light. Many bottom materials have rough surfaces that multiply scatter incident light in the process of eventually reflecting it away from the surface. Therefore, pending actual measurements, it is expedient, and perhaps even reasonable, to model sediments and and aquatic vegetation as Lambertian, depolarizing surfaces. $\underline{r}_{mb}$ then has the simple form

$$\underline{r}_{mb} \equiv \underline{r}_{\mathrm{Lamb}} = \frac{R(\lambda)\cos\theta'}{\pi} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} . \tag{1.45}$$

Here $R(\lambda) = \frac{E_u(\lambda)}{E_d(\lambda)}$ is the wavelength-dependent irradiance reflectance of the bottom for unpolarized light; π carries the units of sr^{-1}. Because only the (1,1) matrix element is nonzero, the rotation matrices, which do not affect the (1,1) matrix element, are not included in Eq. (1.45). Matrices of this form have been used by Zhai et al. (2008) in studies of underwater imaging.

1.6.2 Infinitely Deep Water

The vector reflectance properties of an infinitely deep layer of homogeneous, source-free water below depth $z = m$ are much more complicated than for an opaque bottom and have not yet been worked out. The development would presumably follow that for the scalar case seen in *Light and Water* Section 9.5. The option for an infinitely deep bottom boundary condition can be added after the code has been developed and validated for the case of finite-depth opaque bottoms. Meanwhile, infinitely deep bottoms can be simulated simply by placing an opaque bottom at finite depth m sufficiently deep that the bottom reflectance does not significantly affect the radiance at the greatest

depth of interest. The only penalty for this is increased run time for solving the VRTE to a great depth.

1.7 Summary of the Continuous Mathematical Problem

We can now succinctly state the mathematical problem to be solved. HydroPol solves the one-dimensional, time-independent VRTE

$$\begin{aligned}\mu\theta\frac{d\underline{S}(\zeta,\mu,\phi,\lambda)}{d\zeta} &= -\underline{S}(\zeta,\mu,\phi,\lambda) \\ &\quad + \underset{\sim}{\omega_{\mathrm{o}}(\zeta,\lambda)}\int_0^{2\pi}\int_0^{\pi}\underset{\sim}{\tilde{P}(\zeta;\mu',\phi'\rightarrow\mu,\phi;\lambda)}\,\underline{S}(\zeta,\mu',\phi',\lambda)d\mu'd\phi' \\ &\quad + \underset{\sim}{\underline{\Sigma}(\zeta,\mu,\phi,\lambda)}\ ,\end{aligned} \tag{1.46}$$

between the sea surface at depth $\zeta = 0$ and the bottom at depth $\zeta = m$, subject to the boundary conditions at the sea surface

$$\begin{aligned}\underline{S}(a,\mu,\phi,\lambda) &= \iint_{2\pi_d}\underset{\sim}{\underline{r}_{aw}(\mu',\phi'\rightarrow\mu,\phi;\lambda)}\,\underset{\sim}{\underline{S}(a,\mu',\phi',\lambda)}\,d\mu'd\phi' \\ &\quad + \iint_{2\pi_u}\underset{\sim}{\underline{t}_{wa}(\mu',\phi'\rightarrow\mu,\phi;\lambda)}\,\underline{S}(w,\mu',\phi',\lambda)\,d\Omega(\mu',\phi') \quad \text{for } (\mu,\phi)\in 2\pi_{\mathrm{u}}\end{aligned} \tag{1.47}$$

and

$$\begin{aligned}\underline{S}(w,\mu,\phi,\lambda) &= \iint_{2\pi_u}\underset{\sim}{\underline{r}_{wa}(\mu',\phi'\rightarrow\mu,\phi;\lambda)}\,\underline{S}(w,\mu',\phi',\lambda)\,d\mu'd\phi' \\ &\quad + \iint_{2\pi_d}\underset{\sim}{\underline{t}_{aw}(\mu',\phi'\rightarrow\mu,\phi;\lambda)}\,\underset{\sim}{\underline{S}(a,\mu',\phi',\lambda)}\,d\mu'd\phi' \quad \text{for } (\mu,\phi)\in 2\pi_{\mathrm{d}}\ ,\end{aligned} \tag{1.48}$$

and at the bottom at depth $\zeta = m$

$$\underline{S}(m,\mu,\phi,\lambda) = \iint_{2\pi_d}\underset{\sim}{\underline{r}_{mb}(\mu',\phi'\rightarrow\mu,\phi;\lambda)}\,\underline{S}(m,\mu',\phi',\lambda)\,d\mu'd\phi' \quad \text{for } (\mu,\phi)\in 2\pi_{\mathrm{u}}\ . \tag{1.49}$$

The quantities with a wavy underline are assumed known. These are the the water column IOPs, the internal source (if any), the sky radiance incident onto the air-side of the sea surface, the air-water surface reflectance and transmittance properties, and the bottom reflectance properties. After solution of these equations, the Stokes vector $\underline{S}(\zeta,\theta,\phi,\lambda)$ is known within (depths $0 \le \zeta \le m$) and leaving (depth label $\zeta = a$) the water.

CHAPTER 2

Discretization of the VRTE

The VRTE and its boundary conditions represent a continuum of depths, directions, and wavelengths. In order to solve these equations on a computer, they must be discretized to obtain a finite number of output variables, viz. Stokes vectors for a finite number of depths, directions, and wavelengths. There are various ways in which to do this. For example, a function of direction can be expanded as an infinite series and then truncated after a finite number of terms. The result would then be an approximate Stokes vector for an exact direction. HydroPol, on the other hand, averages functions of direction over finite solid angles. *The resulting Stokes vectors are exact values averaged over finite solid angles.*

2.1 Discretization of Stokes Vectors

Figure 2.1 shows the directional bins used in HydroPol, which are the same as those used in the standard version of HydroLight. Let Ξ represent the set of all directions $\boldsymbol{\xi} = (\theta, \phi)$. Ξ is divided into rectangular regions, called quads, of size 10 deg in polar angle by 15 deg in azimuthal angle, plus two polar caps of 5 deg half angle. The quad boundaries are analogous to lines of constant latitude and longitude on the earth. This unit sphere of directions can be thought of as centered at the origin of the **x**-**y**-**z** coordinate system, which is shown in blue. The red line at the "equator" divides the set of all directions into hemispheres of upward (Ξ_- or Ξ_u) and downward (Ξ_+ or Ξ_d) directions.

As will be seen in Chapter 4, invariant imbedding splits the VRTE into sets of "upward" and "downward" equations. Polar angle θ is normally measured from 0 in the $+\mathbf{z}$ direction to 90 deg at the equator. However, rather than continuing with the measurement of θ to π or 180 deg in the $-\mathbf{z}$ direction, it will be convenient to think of the upward and downward hemispheres as each having θ run from 0 at the pole to 90 deg at the equator. The θ quad boundaries thus run from 5 deg for the polar cap boundary, through 15, 25, ..., 85, and 90 deg, for a total of $M = 10$ $\Delta\theta$ or $\Delta\mu$ bins in each hemisphere. These quads are numbered from $u = 1$ next to the equator to $u = M = 10$ for the polar caps. Because of the partitioning into upward and downward hemispheres, the equator must always be a quad boundary, hence the row of quads next to the equator has $\Delta\theta = 5$ deg. The two 5 deg $\Delta\theta$ quads next to the equator can, if desired, be combined during post-processing of the

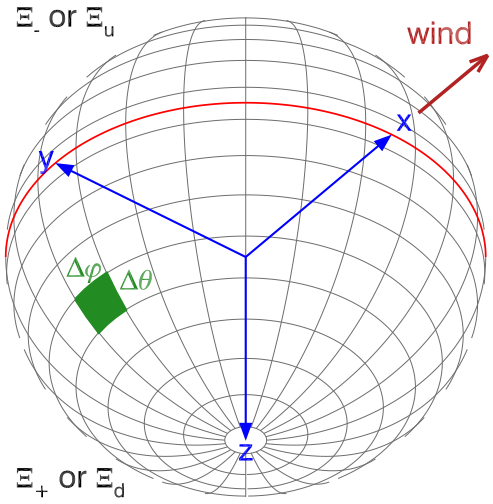

Figure 2.1: The partition of the unit sphere of directions into a finite number of solid angles. The coordinate system is drawn with $+\mathbf{z}$ pointing downward, as in HydroPol. The $+\mathbf{x}$ direction is downwind. The perspective view of this figure can be thought of as being in the water looking obliquely upward.

solution radiances to obtain a horizontal quad of size $\Delta\theta = 10$ deg.

Azimuthal angle ϕ is measured from 0 in the $+\mathbf{x}$ direction, and the first ϕ quad is likewise centered on $+\mathbf{x}$. For reasons that will be seen in §5.1, the coordinate system is "wind centered," with $+\mathbf{x}$ being in the downwind direction. Thus $\mathbf{x}$ is always downwind, $\mathbf{z}$ is downward into the water, and $\mathbf{y} = \mathbf{z} \times \mathbf{x}$ is the crosswind direction. The sun can be placed at any azimuthal direction relative to the wind.

The invariant imbedding solution algorithm used in HydroPol uses a Fourier decomposition in the azimuthal direction, which requires an even number of $\Delta\phi$ bins with each $\Delta\phi$ bin being the same size. We therefore write $2N$ for the total number of azimuthal quads. For 15 deg $\Delta\phi$ quads, there are $2N = 24$ ϕ quads, numbered $v = 1$ to $v = 2N = 24$, with quad 1 being the one centered on $+\mathbf{x}$. In addition, selecting the number of quads to be a multiple of 4, as here, is convenient for calculation of radiances at azimuthal directions of 0, 90, 180, and 270 degrees because those directions are then at the centers of quads.

The total number of quads and polar caps is thus $2(M-1)2N + 2 = 434$. The quad with θ band index u and ϕ band index v is denoted $Q(u,v)$ or Q_{uv}, or $Q(M)$ or $Q_M$for a polar cap. With these conventions, it can be seen that the shaded green quad in Fig. 2.1 corresponds to quad $Q(u,v) = Q(5,7)$ in the downward hemisphere (recall that $+\mathbf{z}$ is downward as emphasized in Fig. 2.1, so the green quad as drawn is in the Ξ_d or Ξ_+ hemisphere). The (θ,ϕ) range of this quad is θ = 45 to 55 deg and ϕ = 82.5 to 97.5 deg. This quad would contain radiance traveling downward at a nominal angle (quad center) of 50 deg from the nadir direction and perpendicular to the wind direction. Polar cap quads have no ϕ dependence and must always be treated as a special case.

The solid angle of $Q(u,v)$ is

$$\begin{aligned}\Omega_{uv} &= \iint_{Q(u,v)} d\Omega = \int_{\Delta\theta_u}\int_{\Delta\phi_v} \sin\theta d\theta d\phi \\ &= \int_{\Delta\mu_u}\int_{\Delta\phi_v} d\mu d\phi = \Delta\mu_u \Delta\phi_v \,, \end{aligned} \tag{2.1}$$

or $\Omega_M = 2\pi[1-\cos(\theta_{(M-1)})] = 2\pi[1-\mu_{(M-1)}]$ for a polar cap, where $\theta_{(M-1)}$ is the boundary of

the polar cap.

To convert the variables in the VRTE from continuous functions of depth, direction, and wavelength, to a finite number of values for discrete depths, directions, and wavelengths, HydroPol

1. Integrates the VRTE over quads of finite solid angle Ω_{uv},
2. Integrates the VRTE over wavelength bands of finite bandwidth $\Delta\lambda_j, j = 1, ..., J$,
3. Saves the solution at only a finite number of depths $z_k, k = 1, ..., K$.

Thus the computed Stokes vectors are

$$\underline{S}(k, u, v, j) = \frac{1}{\Delta\lambda_j} \frac{1}{\Omega_{uv}} \int_{\Delta\lambda_j} \int_{\Delta\mu_u} \int_{\Delta\phi_v} \underline{S}(z_k, \mu, \phi, \lambda) \sin\theta \, d\mu \, d\phi \, d\lambda \,. \tag{2.2}$$

These integrations are applied to each of the four $[I, Q, U, V]^T$ Stokes components. The wavelength bands $\Delta\lambda_j$ do not need to be equal in size.

As will be seen, the depths where output is saved refer to exact depths in the water column, not to finite-depth bins over which variables are depth averaged. HydroPol is not a "layered" model in which the inputs and outputs are assumed to be constant or average values within finite-thickness layers (as is the case, for example, with the discrete ordinates and layer adding and doubling solution methods). Indeed, the IOP inputs to HydroPol must be suppled as continuous functions of depth, and the solution algorithm solves the VRTE with arbitrarily fine depth resolution, which is determined by the accuracy parameters of the numerical differential equation solver. The solution is simply saved at the finite set of user-requested output depths z_k as the VRTE is solved as a continuous function of depth.

The proper interpretation of $\underline{S}(k, u, v, j)$ is that it is the exact Stokes vector $\underline{S}(z, \mu, \phi, \lambda)$ averaged over direction within quad $Q(u, v)$ and over wavelength within band $\Delta\lambda_j$, at a particular depth z_k.

The chosen quad angular partition is a balance between the conflicting needs of having sufficiently high angular resolution in the radiance distribution and keeping the run times acceptably small. Numerical studies with the scalar HydroLight code show that a finer angular resolution does not change computed quantities such as irradiances or AOPs by more than roughly one percent, whereas accuracy starts to degrade for a coarser resolution. However, the run time is proportional the *square* of the number of quads because scattering must be computed from every quad into every other quad. Thus finer angular resolution comes at a high computational cost with almost no improvement in numerical accuracy for the quantities of interest to most users.

The wavelength bands are usually chosen (by the user at run time) to match those of oceanographic sensors, typically $\Delta\lambda = 5$ or 10 nm. Run time is roughly proportional to the number of wave bands. However, the optical depth is wavelength dependent, so wavelengths greater than 700 nm, where water absorption becomes very large, can take much longer than shorter wavelengths.

Neither the number of depths K where output is requested nor their spacing significantly affects run time. Requesting additional output depths does not change the values at the computed depths, it only increase the size of the output files. The size of the output files is directly proportional to the number of output depths, quads, and wavelength bands.

2.2 Quad Averaging

For a general function of direction $F(\theta,\phi) = F(\mu,\phi)$, the average over all directions in quad $Q(u,v)$, termed the quad average $F(u,v)$, is defined as (L&W 4.52)

$$\begin{aligned} F(u,v) &= \frac{1}{\Omega_{uv}} \int_{\Delta\theta_u} \int_{\Delta\phi_v} F(\theta,\phi)\, \sin\theta \, d\theta \, d\phi \\ &= \frac{1}{\Omega_{uv}} \int_{\Delta\mu_u} \int_{\Delta\phi_v} F(\mu,\phi)\, d\mu \, d\phi \,. \end{aligned} \tag{2.3}$$

Note that $F(u,v)$ has the same dimensions as $F(\theta,\phi)$, e.g., units of radiance if F is a Stokes vector.

As an example, let $\mu_u(1), \mu_u(2)$ and $\phi_v(1), \phi_v(2)$ denote the μ and ϕ boundaries of quad $Q(u,v)$. Then the quad average of the polar angle variable μ over this quad is

$$\begin{aligned} \frac{1}{\Omega_{uv}} \int_{\mu_u(1)}^{\mu_u(2)} \int_{\phi_v(1)}^{\phi_v(2)} \mu \, d\mu \, d\phi = & \\ &= \frac{1}{\Omega_{uv}} \frac{1}{2} [\mu_u^2(2) - \mu_u^2(1)][\phi_v(2) - \phi_v(1)] \\ &= \frac{1}{\Omega_{uv}} \frac{1}{2} [\mu_u(2) + \mu_u(1)][\mu_u(2) - \mu_u(1)][\phi_v(2) - \phi_v(1)] \\ &= \frac{1}{\Delta\mu_u \Delta\phi_v} \mu_u \Delta\mu_u \Delta\phi_v \\ &= \mu_u \end{aligned} \tag{2.4}$$

Here $\mu_u = \frac{1}{2}[\mu_u(2) + \mu_u(1)]$ is the average of the quad boundary values. This example belabors a simple calculation, but it does serve to illustrate how quad averaging works.

We can thus see that the quad average of the left-hand side of the VRTE, Eq. (1.22), is

$$\begin{aligned} \frac{1}{\Omega_{uv}} \int_{\Delta\mu_u} \int_{\Delta\phi_v} \mu \frac{d\underline{S}(z,\mu,\phi,\lambda)}{dz} \, d\mu \, d\phi = & \\ &= \frac{d}{dz} \left\{ \frac{1}{\Omega_{uv}} \int_{\Delta\mu_u} \int_{\Delta\phi_v} \mu \underline{S}(z,\mu,\phi,\lambda) \, d\mu \, d\phi \right\} \\ &= \mu_u \frac{d\underline{S}(z,u,v,\lambda)}{dz} \,. \end{aligned} \tag{2.5}$$

Here we have noted that depth and direction are independent variables, so the order of integration over direction and differentiation with respect to depth can be interchanged. This quad-averaged term is discretized in wavelength as shown in Eq. (2.2).

Scattering involves bi-directional functions of incident (μ',ϕ') and scattered (μ,ϕ) directions. The corresponding quad average of a bi-directional function $F(\mu',\phi' \to \mu,\phi)$ is (L&W 4.63)

$$F(r,s \to u,v) = \frac{1}{\Omega_{uv}} \int_{\Delta\mu_u} \int_{\Delta\phi_v} \left[\int_{\Delta\mu_r} \int_{\Delta\phi_s} F(\mu',\phi' \to \mu,\phi) \, d\mu' \, d\phi' \right] d\mu \, d\phi \,. \tag{2.6}$$

The bi-directional functions needed in HydroPol all describe scattering (either in the water column or at a boundary surface) and have units of sr^{-1}. Thus a phase matrix element $P_{i,j}(\psi) = P_{i,j}(\mu',\phi' \to \mu,\phi)$ tells how much radiance is scattered from direction (μ',ϕ') to direction (μ,ϕ) and

from polarization state j to *i per unit solid angle* in the scattered direction. *The corresponding quad-averaged function $P_{i,j}(r, s \to u, v)$ is however non-dimensional.* This is because $P_{i,j}(r, s \to u, v)$ tells how much radiance is scattered in total (not per unit solid angle) from quad $Q(r, s)$ into quad $Q(u, v)$.

As seen in Eq. (1.22), to directionally discretize the VRTE, we must quad average integrals of the form

$$\int_0^{2\pi} \int_0^{\pi} \underline{\tilde{P}}(\mu', \phi' \to \mu, \phi)\, \underline{S}(\mu', \phi') d\mu' d\phi' \,.$$

Quad averaging is applied to each matrix element or product of elements, e.g. $P_{i,j}S_j$, in such quantities. Dropping the matrix and element notation for brevity, the quad average of such an integral becomes

$$\begin{aligned}
&\frac{1}{\Omega_{uv}} \int_{\Delta\mu_u} \int_{\Delta\phi_v} \left\{ \int_0^{2\pi} \int_0^{\pi} P(\mu', \phi' \to \mu, \phi) S(\mu', \phi')\, d\mu'\, d\phi' \right\} d\mu\, d\phi \\
&= \frac{1}{\Omega_{uv}} \int_{\Delta\mu_u} \int_{\Delta\phi_v} \left\{ \sum_r \sum_s \int_{\Delta\mu_r} \int_{\Delta\phi_s} P(\mu', \phi' \to \mu, \phi) S(\mu', \phi')\, d\mu'\, d\phi' \right\} d\mu\, d\phi \\
&= \sum_r \sum_s \left\{ \frac{1}{\Omega_{uv}} \int_{\Delta\mu_u} \int_{\Delta\phi_v} \left[\int_{\Delta\mu_r} \int_{\Delta\phi_s} P(\mu', \phi' \to \mu, \phi) S(\mu', \phi')\, d\mu'\, d\phi' \right] d\mu\, d\phi \right\} \\
&= \sum_r \sum_s P(r, s \to u, v) S(r, s) \,. \qquad (2.7)
\end{aligned}$$

In the second equation, the integral over all incident directions has been written as a sum of integrals over all quads $Q(r, s)$. The next equation interchanges the order of integration over the scattered-direction quad $Q(u, v)$ and the summation over incident direction quads. The final equations uses the definition of Eq. (2.6). Note that the quad average of the integrals reduces to sums over the quad-averaged integrand. The solid angle factors resulting from the differentials $d\mu\, d\phi$ are built into the quad-averaged terms; there are thus no explicit $\Delta\mu_u \Delta\phi_v$ factors in the final expression.

These results are sufficient to show that the directional and wavelength discretized VRTE is

$$\begin{aligned}
\mu_u \frac{d\underline{S}(\zeta, u, v, j)}{d\zeta} \;&= -\, \underline{S}(\zeta, u, v, j) \\
&\quad + \omega_o(\zeta, j) \sum_r \sum_s \underline{\tilde{P}}(\zeta; r, s \to u, v; j)\, \underline{S}(\zeta, r, s, j) \\
&\quad + \underline{\Sigma}(\zeta, u, v, j) \,. \qquad (2.8)
\end{aligned}$$

Symbolically, $u = 1, ..., 2M$ and $v = 1, ..., 2N$ to cover all directions. The polar cap special cases will be considered below. The j index represents the j^{th} wavelength band. This is the system of equations solved by HydroPol, subject to discretized boundary conditions at the sea surface and bottom.

It should be noted that the development of this discretized VRTE via quad averaging did not make any assumptions about the phase matrix (or phase function, in the scalar theory) other than isotropy of the scattering medium. In particular, it was not necessary to expand the phase matrix as an infinite series of Legendre polynomials, and then truncate that series to a finite number of terms, as is required by the discrete ordinates solution method (e.g., L&W 9.1). Such expansions cause numerical difficulties for highly peaked phase functions (e.g., L&W Fig. 9.1), which are typical of

ocean waters. *One of the most important conceptual and numerical virtues of quad averaging is that it can handle arbitrary phase functions.*

Discretization of the surface boundary conditions (1.40) and (1.41) follows in a similar fashion to that for the path integral in the VRTE. The result is

$$\begin{aligned}\underline{S}(a,u,v) =& \sum_{\substack{r \\ r,s\in\Xi_d}} \sum_{s} \underline{r}_{aw}(r,s \to u,v)\, \underline{S}(a,r,s) \\ &+ \sum_{\substack{r \\ r,s\in\Xi_u}} \sum_{s} \underline{t}_{wa}(r,s \to u,v)\, \underline{S}(w,r,s) \quad \text{for (u, v)} \in \Xi_{\mathrm{u}}\end{aligned} \tag{2.9}$$

and

$$\begin{aligned}\underline{S}(w,u,v) =& \sum_{\substack{r \\ r,s\in\Xi_u}} \sum_{s} \underline{r}_{wa}(r,s \to u,v)\, \underline{S}(w,r,s) \\ &+ \sum_{\substack{r \\ r,s\in\Xi_d}} \sum_{s} \underline{t}_{aw}(r,s \to u,v)\, \underline{S}(a,r,s) \quad \text{for (u, v)} \in 2\Xi_{\mathrm{d}}\,.\end{aligned} \tag{2.10}$$

The discretized transfer functions are computed as in (2.6), e.g.,

$$\underline{t}_{aw}(r,s \to u,v) = \frac{1}{\Omega_{uv}} \int_{\Delta\mu_u} \int_{\Delta\phi_v} \left[\int_{\Delta\mu_r} \int_{\Delta\phi_s} \underline{t}_{aw}(\mu',\phi' \to \mu,\phi)\, d\mu'\, d\phi' \right] d\mu\, d\phi\,. \tag{2.11}$$

Note that, just as in the scalar case, the discretized function $\underline{t}_{aw}(r,s \to u,v)$ is nondimensional, whereas the continuous function of direction, $\underline{t}_{aw}(\mu',\phi' \to \mu,\phi)$, has units of sr^{-1}. Finally, the discretized bottom boundary condition is

$$\underline{S}(b,u,v) = \sum_{\substack{r \\ r,s\in\Xi_d}} \sum_{s} \underline{r}_{b}(r,s \to u,v)\, \underline{S}(b,r,s) \quad \text{for (u, v)} \in \Xi_{\mathrm{u}}. \tag{2.12}$$

Finally, note that polar caps are always a special case because they have no ϕ dependence. Thus the sums represented symbolically by $\sum_r \sum_s$ must explicitly account for the polar caps $r = M$ and s undefined. In the computer code, the polar cap values are stored in array location $(r,s) = (M,1)$, with the array elements $(r = M, s = 2, ..., 2N)$ being unused (and set to zero for convenience). Thus the path radiance term for scattering into $Q(u,v)$ is evaluated in the computer code as

$$\begin{aligned}&\sum_{r} \sum_{s} \underline{\tilde{P}}(\zeta; r,s \to u,v; j)\, \underline{S}(\zeta,r,s,j) = \\ &\sum_{r=1}^{M-1} \sum_{s=1}^{2N} \underline{\tilde{P}}(\zeta; r,s \to u,v; j)\, \underline{S}(\zeta,r,s,j) + \underline{\tilde{P}}(\zeta; M,1 \to u,v; j)\, \underline{S}(\zeta,M,1,j)\,.\end{aligned} \tag{2.13}$$

2.2.1 Numerical Evaluation of Quad Averages

Quad averages of functions $F(\theta,\phi)$ of direction, e.g., the four elements of the sky radiance $\underline{S}(a,\theta,\phi)$, are defined by Eq. (2.3). Quad averages of bi-directional functions are defined by Eq. (2.6). There are two ways to compute quad averages. If the function is an analytic function, numerical integration can be used. This is the case for phase matrices, or for the Fresnel reflectance and transmittance functions for a level sea surface. However, if the function represents a surface reflectance or transmittance for a wind-blown sea surface, no analytic function exists, and Monte Carlo simulation must be used. That situation will be described in §3.2.

Quad averages of analytic functions $F(\theta,\phi)$ are numerically computed as a double summation over subquads of each quad $Q(u,v)$. A $Q(u,v)$ quad is divided into $n_\mu \ge 1$ and $n_\phi \ge 1$ subquads of size $\delta_\mu(u) = \Delta\mu_u/n_\mu$, $\delta_\phi(v) = \Delta\phi_v/n_\phi$, etc. The double integral of Eq. (2.3) is then approximated as a double summation over the sub-quads, e.g.

$$F(u,v) = \frac{1}{\Omega_{uv}}\delta_\mu(u)\delta_\phi(v)\sum_{i=1}^{n_\mu}\sum_{j=1}^{n_\phi} F(\mu_i(u),\phi_j(v))\,. \tag{2.14}$$

Phase matrices, surface reflectance and transmission functions for a level sea surface, and bottom VBRDF functions involve two directions as seen in Eq. (2.6). Quad averages (and the associated Fourier amplitudes) of these functions are pre-computed and stored for repeated use in solving the VRTE for various IOP and other inputs. A quadruple integration like that shown in Eq. (2.11) is computed as follows. The $Q(r,s)$ and $Q(u,v)$ quads are divided into $n_\mu \ge 1$ and $n_\phi \ge 1$ subquads of just as for Eq. (2.14). The integrals of Eq. (2.6) are then approximated as a quadruple summation over the sub-quads (L&W 8.13):

$$F(r,s\to u,v) = \frac{1}{\Omega_{uv}}\delta_\mu(u)\delta_\phi(v)\delta_\mu(r)\delta_\phi(s)\sum_{i=1}^{n_\mu}\sum_{j=1}^{n_\phi}\sum_{k=1}^{n_\mu}\sum_{\ell=1}^{n_\phi} F(\psi_{ijk\ell})\,. \tag{2.15}$$

where $\psi_{ijk\ell}$ is the angle between the centers of the subquads. $\psi_{ijk\ell}$ is obtained from Eq. (1.4):

$$\cos(\psi_{ijk\ell}) = \mu_i(u)\mu_k(r) + \sqrt{1-\mu_i^2(u)}\sqrt{1-\mu_k^2(r)}\cos[\phi_j(v)-\phi_\ell(s)]\,. \tag{2.16}$$

Experience with the discretization of phase functions in HydroLight shows that $n_\mu = 3$ and $n_\phi = 4$, which give roughly 3×4 deg subquads, work well except for quads with near-forward scattering and highly peaked phase functions. For adjacent quads, which correspond to small scattering angles, the values of n_μ and n_ϕ are increased by a factor of 10 to get high angular resolution of phase functions at small scattering angles. A similar increase in subquad resolution can be done in Eq. (2.14) if needed, e.g. for the computation of quad averaged sky radiance for the quad containing the sun, for which the radiance depends strongly on direction near the sun's location.

Finally, when the incident and scattered quads are the same, so that the scattering angle can be very small, the forward-scatter-quad phase function values are corrected via the normalization condition (1.28). That equation is discretized by writing the integration over all (θ,ϕ) as a sum of integrations over all quads $Q(u,v)$, integrating the equation over $Q(r,s)$, and applying the definition of Eq. (2.6). The resulting discretized version of Eq. (1.28) is

$$\frac{1}{\Omega_{rs}}\sum_u\sum_v \tilde{\beta}(r,s\to u,v)\Omega_{uv} = 1\,. \tag{2.17}$$

That is, for each incident quad $Q(r,s)$, the sum of $\tilde{\beta}(r,s \to u,v)\Omega_{uv}/\Omega_{rs}$ over all scattered quads $Q(u,v)$ is 1. After discretization, this sum is evaluated as a check on the accuracy of the phase function numerical discretization of Eq. (2.15). The sum over quads is usually within a few percent of the correct value of 1 for all (r,s) values. Phase functions change most rapidly at very small scattering angles, and are also then the hardest to measure. Thus most of the error in discretizing phase functions occurs for the forward scatter quad, i.e. when $(r,s) = (u,v)$. Thus a final correction to the forward-scatter quads is made by solving this equation for the forward-scatter term (L&W 8.14):

$$\tilde{\beta}(r,s \to r,s) = 1 - \frac{1}{\Omega_{rs}} \sum_{\substack{u \\ (u,v)\neq(r,s)}} \sum_{v} \tilde{\beta}(r,s \to u,v)\Omega_{uv} \,. \tag{2.18}$$

Equation (2.18) is then used to recompute the forward-scatter quads. This forces the discretized phase function to exactly satisfy the required normalization condition (2.17) regardless of any errors in the evaluation of Eq. (2.15). Since the phase function is just the (1,1) element of the phase matrix, this same renormalization is used in HydroPol for the (1,1) element of the phase matrix:

$$\tilde{P}_{11}(r,s \to r,s) = 1 - \frac{1}{\Omega_{rs}} \sum_{\substack{u \\ (u,v)\neq(r,s)}} \sum_{v} \tilde{P}_{11}(r,s \to u,v)\Omega_{uv} \,. \tag{2.19}$$

Unfortunately, there does not seem to be a corresponding check for the other elements of the phase matrix. However, the other elements of the reduced phase matrix are well behaved, so if the numerical integration for the (1,1) element is accurate, then the discretization for the other elements should also be accurate.

2.3 Summary of the Discretized Mathematical Problem

We can now summarize the discretized mathematical problem. HydroPol solves the one-dimensional, time-independent discretized VRTE

$$\begin{aligned} \mu_u \frac{d\underline{S}(\zeta,u,v,j)}{d\zeta} &= -\underline{S}(\zeta,u,v,j) \\ &\quad + \underset{\sim}{\omega_o(\zeta,j)} \sum_r \sum_s \underset{\sim}{\tilde{P}(\zeta;r,s \to u,v;j)}\, \underline{S}(\zeta,r,s,j) \\ &\quad + \underset{\sim}{\underline{\Sigma}(\zeta,u,v,j)} \,, \end{aligned} \tag{2.20}$$

between the sea surface at depth $\zeta = 0$ and the bottom at depth $\zeta = b$, subject to the boundary conditions at the sea surface

$$\begin{aligned} \underline{S}(a,u,v,j) = &\sum_{\substack{r \\ r,s\in\Xi_d}} \sum_{s} \underset{\sim}{\underline{r}_{aw}(r,s \to u,v;j)}\, \underset{\sim}{\underline{S}(a,r,s,j)} \\ &+ \sum_{\substack{r \\ r,s\in\Xi_u}} \sum_{s} \underset{\sim}{\underline{t}_{wa}(r,s \to u,v;j)}\, \underline{S}(w,r,s,j) \quad \text{for (u, v)} \in \Xi_{\text{u}} \end{aligned} \tag{2.21}$$

and

$$\begin{aligned}\underline{S}(w,u,v,j) =& \sum_{\substack{r \\ r,s\in\Xi_u}} \sum_{s} \underset{\sim}{\underline{r}_{wa}}(r,s\rightarrow u,v;j)\,\underline{S}(w,r,s,j) \\ &+\sum_{\substack{r \\ r,s\in\Xi_d}} \sum_{s} \underset{\sim}{\underline{t}_{aw}}(r,s\rightarrow u,v;j)\,\underline{\underset{\sim}{S}}(a,r,s,j) \quad \text{for } (\mathrm{u},\mathrm{v})\in\Xi_{\mathrm{d}}\ , \end{aligned} \tag{2.22}$$

and at the bottom at depth $\zeta = b$

$$\underline{S}(m,u,v,j) = \sum_{\substack{r \\ r,s\in\Xi_d}} \sum_{s} \underset{\sim}{\underline{r}_{mb}}(r,s\rightarrow u,v;j)\,\underline{S}(m,r,s,j) \quad \text{for } (\mathrm{u},\mathrm{v})\in\Xi_{\mathrm{u}}\ . \tag{2.23}$$

The quantities with a wavy underline are assumed known. These are the the water column IOPs, the internal source (if any), the sky radiance incident onto the air-side of the sea surface, the air-water surface reflectance and transmittance properties, and the bottom reflectance properties, all in discretized form. After solution of these equations, the Stokes vector $\underline{S}(\zeta,u,v,j)$ is known within (depths $0\le\zeta\le m$) and leaving (depth label $\zeta = a$) the water.

CHAPTER 3

The Air-Water Surface

The air-water surface boundary conditions seen in Eqs. (2.21) and (2.21) are stated in terms of four surface transfer functions, $\underline{r}_{aw}$, $\underline{t}_{aw}$, $\underline{r}_{wa}$, and $\underline{t}_{wa}$, which are assumed known. These functions are the inherent optical properties of the sea surface. They depend on the surface wave state and water index of refraction, but not on the light incident onto the surface from above or below. This chapter shows how to compute these air-water surface transfer functions. We first consider the case of a level sea surface, and then the much more complicated situation of a wind-blown, wave-covered surface.

3.1 Reflection and Transmission by a Level Sea Surface

Consider a level or flat air-water surface. This simplest of environmental conditions is important for three reasons. First, a wind-blown sea surface can be modeled as a collection of randomly tilted, but locally flat, wave facets. Each tilted facet reflects and transmits light according to the laws for a flat surface. Thus a full understanding of how a flat surface reflects and transmits light is the foundation of modeling wind-blown surfaces. Second, the equations for a flat surface are analytically tractable and provide an important check on the numerical computations in the limit of zero wind speed. Third, level sea surfaces occasionally do occur in nature in the absence of wind-generated waves or swell.

Light can be incident onto this surface from the air, in which case part is reflected back to the air by the surface and part is transmitted through the surface into the water. Light can also be incident onto the underside of the sea surface, in which case part (or all) is reflected back to the water and part can be transmitted through the surface into the air. These two situations are illustrated in Fig. (3.1). For either air- or water-incident light, $\underline{S}_i$ denotes the Stokes vector of the incident light, $\underline{S}_r$ is the reflected light, and $\underline{S}_t$ is the transmitted light. Angles θ_i, θ_r, and θ_t are the incident, reflected, and transmitted directions of the light propagation measured relative to the normal to the surface. $\underline{S}_i$, $\underline{S}_r$, and $\underline{S}_t$ all lie in the same plane.

Just as for scattering within the water, the "surface scattering" processes of reflection and refraction are described by 4×4 scattering matrices. There are four of these matrices: $\underline{R}_{aw}$ describes how air-incident light is reflected by the water surface back to the air, $\underline{T}_{aw}$ describes

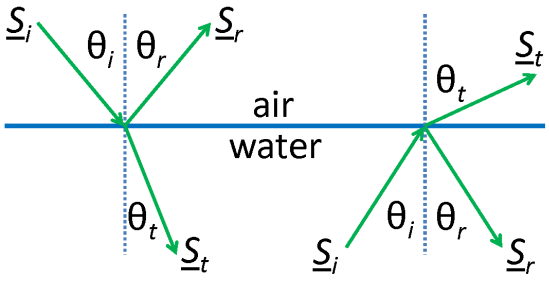

Figure 3.1: Reflection and transmission by a level air-water surface.

how air-incident light is transmitted through the surface into the water, $\underline{R}_{wa}$ reflects water-incident light back to the water, and $\underline{T}_{wa}$ transmits light from the water into the air. However, because $\underline{S}_i$, $\underline{S}_r$, and $\underline{S}_t$ are coplanar, scattering by the level surface does not involve rotation matrices as does scattering within the water body. (Or, from another viewpoint, the incident and final meridian planes and the scattering plane are all the same, the rotation angles α' and α are thus both 0, and the rotation matrices reduce to identity matrices.)

The reflection and (especially) transmission of polarized light by a dielectric surface such as a level water surface are rather complicated processes, and the literature contains a number of different (and, indeed, sometimes incorrect) mathematical formulations of the equations. The formulas given in Garcia (2012) are appropriate to the needs of HydroPol and are used here. Note, however, that although the equations in Garcia (2012) are correct, some of his derivations and interpretations are incorrect, as explained by Zhai et al. (2012). Both papers must be used to understand the equations now presented. The equations in Garcia (2012) will be referenced by (G21) and so on; the corresponding equations in Zhai et al. (2012) will be referenced as (Z5), etc.

The reflectance and transmittance matrices have a general formulation for the interface between any two dielectric media a and b. Let n_a be the index of refraction of medium a and n_b be that of medium b. In general n_a and n_b are complex numbers, but for the air-water surface we take $n_{\text{air}} = 1$ and $n_{\text{water}} \approx 1.34$ to be the real indices of refraction. For reflection, the reflected angle θ_r equals the incident angle θ_i. For transmission from a to b, the transmitted angle is given by Snel's law, $n_a \sin\theta_a = n_b \sin\theta_b$, or

$$\theta_b = \sin^{-1}\left(\frac{n_a \sin\theta_i}{n_b}\right) . \tag{3.1}$$

For water-incident light, $n_a = n_{\text{water}}$ and $n_b = n_{\text{air}}$, in which case the transmitted angle becomes undefined beyond the *critical angle for total internal reflection*, which for water is $\theta_{\text{tir}} = \sin^{-1}(1/n_{\text{water}}) \approx 48\,\text{deg}$. For water-incident angles greater than θ_{tir} the incident light is totally reflected back to the water and no light is transmitted to the air.

Let $\underline{R}_{ab}$ denote the reflectance matrix for radiant *energy* (or power) incident from medium a and reflected back by medium b. $\underline{R}_{ab}$ thus represents either $\underline{R}_{aw}$ or $\underline{R}_{wa}$. Likewise, let $\underline{T}_{ab}$ denote the transmittance matrix for light incident from medium a and transmitted through the surface into medium b. $\underline{T}_{ab}$ thus represents either $\underline{T}_{aw}$ or $\underline{T}_{wa}$.

With these preliminaries, the reflectance matrix $\underline{R}_{ab}$ is (G10)

$$\underline{R}_{ab} = \begin{bmatrix} \frac{1}{2}(R_\parallel R_\parallel^* + R_\perp R_\perp^*) & \frac{1}{2}(R_\parallel R_\parallel^* - R_\perp R_\perp^*) & 0 & 0 \\ \frac{1}{2}(R_\parallel R_\parallel^* - R_\perp R_\perp^*) & \frac{1}{2}(R_\parallel R_\parallel^* + R_\perp R_\perp^*) & 0 & 0 \\ 0 & 0 & \Re\{R_\parallel R_\perp^*\} & \Im\{R_\parallel R_\perp^*\} \\ 0 & 0 & -\Im\{R_\parallel R_\perp^*\} & \Re\{R_\parallel R_\perp^*\} \end{bmatrix} . \tag{3.2}$$

Here $\Re\{R_\parallel R_\perp^*\}$ denotes the real part of $R_\parallel R_\perp^*$ and $\Im\{R_\parallel R_\perp^*\}$ is the imaginary part.

The transmission matrix $\underline{T}_{ab}$ is (G11 or Z3)

$$\underline{T}_{ab} = f_T \begin{bmatrix} \frac{1}{2}(T_\parallel T_\parallel^* + T_\perp T_\perp^*) & \frac{1}{2}(T_\parallel T_\parallel^* - T_\perp T_\perp^*) & 0 & 0 \\ \frac{1}{2}(T_\parallel T_\parallel^* - T_\perp T_\perp^*) & \frac{1}{2}(T_\parallel T_\parallel^* + T_\perp T_\perp^*) & 0 & 0 \\ 0 & 0 & \Re\{T_\parallel T_\perp^*\} & \Im\{T_\parallel T_\perp^*\} \\ 0 & 0 & -\Im\{T_\parallel T_\perp^*\} & \Re\{T_\parallel T_\perp^*\} \end{bmatrix} . \tag{3.3}$$

The components of these equations are given by (G7):

$$R_\parallel = \frac{n_b \cos\theta_a - n_a \cos\theta_b}{n_b \cos\theta_a + n_a \cos\theta_b} \tag{3.4}$$

$$R_\perp = \frac{n_a \cos\theta_a - n_b \cos\theta_b}{n_a \cos\theta_a + n_b \cos\theta_b} \tag{3.5}$$

$$T_\parallel = \frac{2n_a \cos\theta_a}{n_b \cos\theta_a + n_a \cos\theta_b} \tag{3.6}$$

$$T_\perp = \frac{2n_a \cos\theta_a}{n_a \cos\theta_a + n_b \cos\theta_b} . \tag{3.7}$$

In general, the indices of refraction are complex numbers and these equations must be used. However, for real indices of refraction, as is the case in HydroPol, the matrix elements can be simplified at the expense of having a special case for water-incident angles greater that the critical angle.

Define

$$n_{ab} = \frac{n_a}{n_b} \quad \text{and} \quad n_{ba} = \frac{n_b}{n_a} . \tag{3.8}$$

Then for the case of air-incident light, i.e., $n_a \le n_b$, or water-incident light with the incident angle less than the critical angle, i.e., $n_a > n_b$ and $\theta_a < \theta_{\text{tir}}$, the equations yield the real forms (G14 and

G15)

$$R_\parallel R_\parallel^* = \left(\frac{\cos\theta_a - n_{ab}\cos\theta_b}{\cos\theta_a + n_{ab}\cos\theta_b}\right)^2 \tag{3.9}$$

$$R_\perp R_\perp^* = \left(\frac{n_{ab}\cos\theta_a - \cos\theta_b}{n_{ab}\cos\theta_a + \cos\theta_b}\right)^2 \tag{3.10}$$

$$\Re\{R_\parallel R_\perp^*\} = \left(\frac{\cos\theta_a - n_{ab}\cos\theta_b}{\cos\theta_a + n_{ab}\cos\theta_b}\right)\left(\frac{n_{ab}\cos\theta_a - \cos\theta_b}{n_{ab}\cos\theta_a + \cos\theta_b}\right) \tag{3.11}$$

$$\Im\{R_\parallel R_\perp^*\} = 0 \tag{3.12}$$

$$T_\parallel T_\parallel^* = \left(\frac{2n_{ab}\cos\theta_a}{\cos\theta_a + n_{ab}\cos\theta_b}\right)^2 \tag{3.13}$$

$$T_\perp T_\perp^* = \left(\frac{2n_{ab}\cos\theta_a}{n_{ab}\cos\theta_a + \cos\theta_b}\right)^2 \tag{3.14}$$

$$\Re\{T_\parallel T_\perp^*\} = \frac{4n_{ab}^2\cos^2\theta_a}{(\cos\theta_a + n_{ab}\cos\theta_b)(n_{ab}\cos\theta_a + \cos\theta_b)} \tag{3.15}$$

$$\Im\{T_\parallel T_\perp^*\} = 0\,. \tag{3.16}$$

It should be noted that for the case of normal incidence, $\theta_i = 0$, both $R_\parallel R_\parallel^*$ and $R_\perp R_\perp^*$ reduce to

$$R_\parallel R_\parallel^* = R_\perp R_\perp^* = \left(\frac{n_b - n_a}{n_b - n_a}\right)^2\,. \tag{3.17}$$

This gives a reflectance of $R_{ab}(\theta_i = 0) = 0.0209$ for $n_{\text{water}} = 1.338$, for both air- and water-incident light.

For the case of total internal reflection, i.e., $n_a > n_b$ and $\theta_a \geq \theta_{\text{tir}}$, the following equations are to be used (G22):

$$R_\parallel R_\parallel^* = 1 \tag{3.18}$$

$$R_\perp R_\perp^* = 1 \tag{3.19}$$

$$\Re\{R_\parallel R_\perp^*\} = \frac{2\sin^4\theta_a}{1 - (1 + n_{ab}^2)\cos^2\theta_a} - 1 \tag{3.20}$$

$$\Im\{R_\parallel R_\perp^*\} = -\frac{2\cos\theta_a \sin^2\theta_a\sqrt{\sin^2\theta_a - n_{ba}^2}}{1 - (1 + n_{ab}^2)\cos^2\theta_a} \tag{3.21}$$

and all elements of the transmission matrix elements are 0:

$$\underline{T}_{ab} = \underline{T}_{wa} = \underline{0}_{4\times4}\,, \tag{3.22}$$

where $\underline{0}_{4\times4}$ is the 4×4 matrix of zeros.

Finally, the all-important transmission factor f_T in Eq. (3.3) requires discussion. The f_T factor is derived from conservation of energy across the surface. However, that conservation law can be expressed in terms of either radiance or irradiance. For coherent Stokes vectors with units of irradiance, f_T is given by (Zhai et al. (2012), Eq. 9, or Hecht (1989), Eq. 4.59)

$$f_T = \frac{n_b\cos\theta_b}{n_a\cos\theta_a}\,. \tag{3.23}$$

The cosine factors result from the different cross-sectional areas of the incident and transmitted beams due to the change in direction given by Snel's law. The index-of-refraction factors result from the different rates at which energy is transported toward or away from the surface in the two media, i.e., from the differences in the speeds of light in the two media. For diffuse Stokes vectors with units of radiance, as is the case in the present discussion, f_T is given by (Zhai et al. (2012), Eq. 5)

$$f_T = n_{ba}^2 \left(\frac{n_b \cos\theta_b}{n_a \cos\theta_a} \right) . \tag{3.24}$$

The extra factor of n_{ba}^2 in the radiance version results from the change in solid angle when crossing the surface, i.e., for the n^2 law for radiance. The f_T value given in Eq. (3.23) is the form to be used for Monte Carlo ray tracing of individual photons, which carry energy but do not have an associated solid angle. In that case, the n^2 law for radiance is built in to the radiance reflectance functions photon by photon as their directions change when crossing the interface according to Snel's law.

The preceding equations give everything needed to compute the reflection and transmission of light by a level sea surface, either by analytical formulas in the case of a level surface, or by Monte Carlo simulation in the case of a wind-blown sea surface.

As is usually the case, the literature contains a variety of notations, terminology, and formulations. The $R_\parallel, R_\perp, T_\parallel, T_\perp$ terms seen above specify how electric field *amplitudes* are reflected and transmitted for electric fields parallel and perpendicular to the dielectric interface. Snel's law can be used to obtain forms equivalent to Eqs. (3.4)-(3.7) in which the indices of refraction do not appear explicitly (e.g., Born and Wolf (1975), Eqs. 20a, 21a):

$$R_\parallel = \frac{\tan(\theta_a - \theta_b)}{\tan(\theta_a + \theta_b)} \tag{3.25}$$

$$R_\perp = -\frac{\sin(\theta_a - \theta_b)}{\sin(\theta_a + \theta_b)} \tag{3.26}$$

$$T_\parallel = \frac{2\sin\theta_b \cos\theta_a}{\sin(\theta_a + \theta_b)\cos(\theta_a - \theta_b)} \tag{3.27}$$

$$T_\perp = \frac{2\sin\theta_b \cos\theta_a}{\sin(\theta_a + \theta_b)} . \tag{3.28}$$

The fractions of reflected and transmitted *energy* are obtained from the square of the electric field amplitudes. Hence these terms appear as products in Eqs. (3.2) and (3.2). The (1,1) element of $\underline{R}_{ab}$ gives the irradiance (energy) reflectance for air-incident unpolarized light. Using the alternate forms just above gives the Fresnel reflectance for unpolarized irradiance seen in Light and Water, Eq. (4.14a):

$$R_F = \frac{1}{2}(R_\parallel R_\parallel^* + R_\perp R_\perp^*) = (\frac{1}{2}\left\{ \left[\frac{\sin(\theta_a - \theta_b)}{\sin(\theta_a + \theta_b)} \right]^2 + \left[\frac{\tan(\theta_a - \theta_b)}{\tan(\theta_a + \theta_b)} \right]^2 \right\} . \tag{3.29}$$

The corresponding transmitted irradiance for unpolarized light is

$$\begin{aligned} T_F &= f_T \frac{1}{2}(T_\parallel T_\parallel^* + T_\perp T_\perp^*) \\ &= \frac{n_b \cos\theta_b}{n_a \cos\theta_a} \frac{1}{2} \left\{ \left[\frac{2\sin\theta_b \cos\theta_a}{\sin(\theta_a + \theta_b)\cos(\theta_a - \theta_b)} \right]^2 + \left[\frac{2\sin\theta_b \cos\theta_a}{\sin(\theta_a + \theta_b)} \right]^2 \right\} = 1 - R_F , \end{aligned} \tag{3.30}$$

where the f_T of Eq. (3.23) is used for energy transfer.

Figure 3.2 shows the $\underline{R}_{aw}$ and $\underline{T}_{aw}$ matrices as a function of incident angle θ_i for $n_{\rm air} = 1$ and $n_{\rm water} = 1.338$. The (1,1) matrix elements are shown in the upper-left plot, and the (4,4) elements are in the lower-right plot. The red curves are $\underline{R}_{aw}(\theta_i)$ and the blue curves are $\underline{T}_{aw}(\theta_i)$. The reflectance curve for $R_{aw}(1,1)$ is the same as seen in any elementary physics text (e.g. Eq. 3.29): it starts at 0.0209 for normal incidence (as shown above for $n_{\rm water} = 1.338$) and rises to 1 at grazing incidence. The transmission curve for $T_{aw}(1,1)$ on the other hand may look incorrect because it has values greater than one. Its value at normal incidence is

$$T_{aw}(1,1) = \frac{4n_b^3}{(1+n_b)^2} = 1.7528$$

However, this value is indeed correct and is a consequence of the fact that we are dealing with a diffuse Stokes vector with units of radiance. When radiance travels from air to water, the solid angle in air is reduced by a factor of $1/n_{\rm water}^2$ (a consequence of Snel's law), which increases the radiance by a factor of $n_{\rm water}^2$ compared to the radiance for the same amount of power propagating in air. This is termed the *n-squared law for radiance*. The curves in Fig. (3.2) agree exactly with the corresponding plots in Garcia (2012, his Figs. 1-3).

If we were dealing with coherent Stokes vectors with units of irradiance, then the f_T factor of Eq. (3.23) would be used.. The transmittance for normal incidence then would then be

$$T_{aw}(1,1) = (4n_b)/(1+n_b)^2 = 0.9791\,,$$

which with the reflectance sums to one (and also sums to one for all other incident angles). It is, after all, the Law of Conservation of Energy, not the law of conservation of radiance.

The vertical dotted line in Fig. (3.2) shows the location of Brewster's angle,

$$\theta_{Brew} = \tan^{-1}(n_b)\,, \tag{3.31}$$

which is $\tan^{-1}(1.338) = 53.23\,\text{deg}$ in the present case. At this angle, $R_{aw}(1,2) = R_{aw}(2,1) = -R_{aw}(1,1)$, and $R_{aw}(3,3) = R_{aw}(4,4) = 0$. In the present case $R_{aw}(1,1) \approx 0.04$ at θ_{Brew}, and the reflection process $\underline{S}_r = \underline{R}_{aw}(\theta_i = \theta_{Brew})\underline{S}_i$ becomes

$$\underline{S}_r = \begin{bmatrix} 0.04 & -0.04 & 0 & 0 \\ -0.04 & 0.04 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} I \\ 0 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 0.04I \\ -0.04I \\ 0 \\ 0 \end{bmatrix}\,. \tag{3.32}$$

With our choice of "parallel" and "perpendicular" referring to the meridian plane, this R_{aw} has the form of a horizontal polarizer, as was discussed in connection with Eq. (1.17), and $\underline{S}_r = [0.04I, -0.04I, 0, 0]^T$ represents light linearly polarized in a direction parallel to the mean sea surface. Thus, at the Brewster's angle, unpolarized incident radiance is totally horizontally polarized upon reflection.

It should also be noted that the non-zero $T_{aw}(2,1)$ means that unpolarized radiance becomes partly linearly polarized upon transmission through the surface. Thus transmission by the air-water surface induces a partially polarized underwater light field, even if the sea surface is illuminated by unpolarized light.

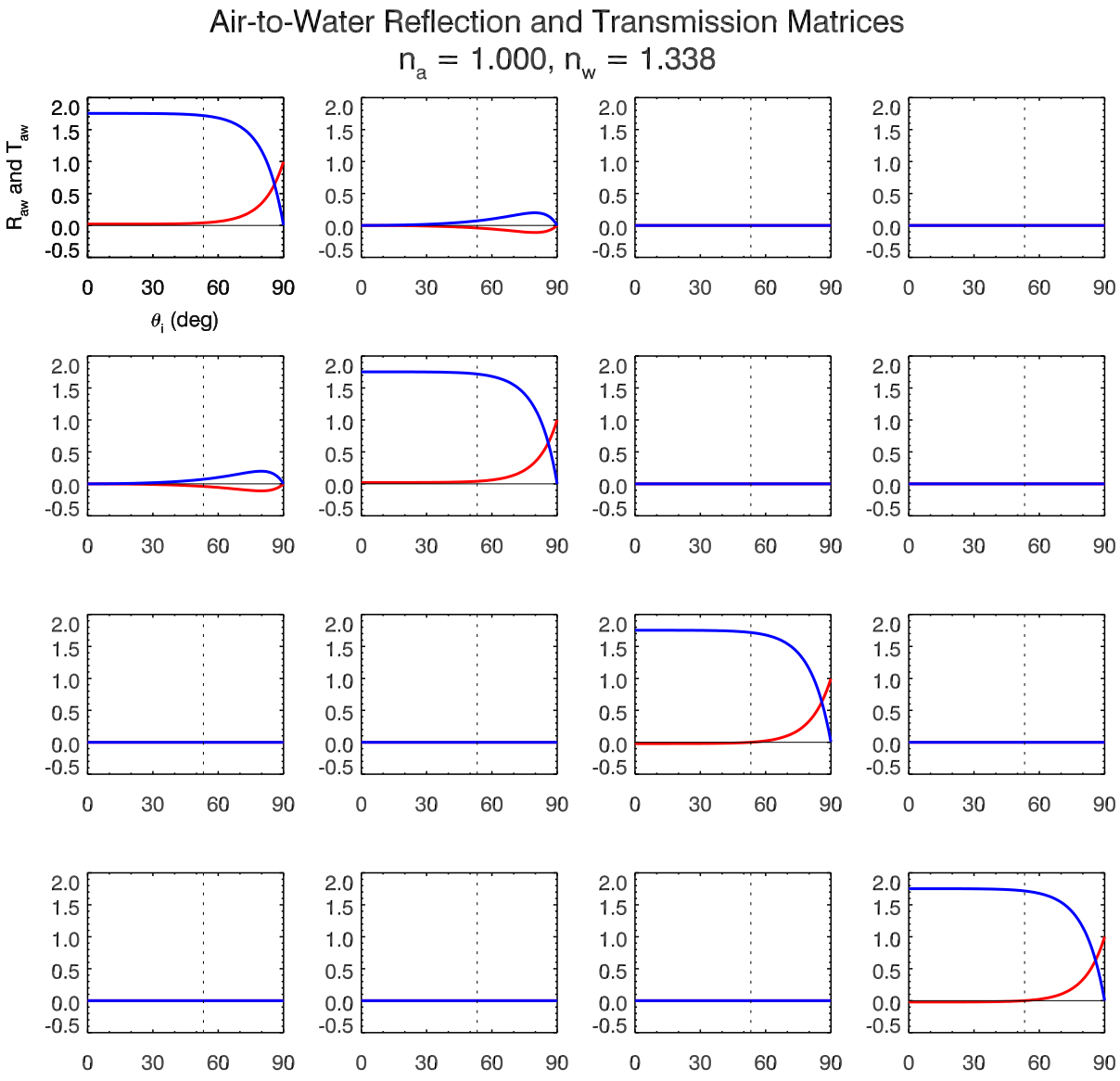

Figure 3.2: Reflectance and transmittance matrices as functions of the incident angle θ_i for air-incident radiance. $\underline{R}_{aw}$ is in red and $\underline{T}_{aw}$ is in blue. The vertical dotted line at $\theta_i = 53.23\,\text{deg}$ is Brewster's angle.

Figure (3.3) shows $\underline{R}_{aw}$ and $\underline{T}_{aw}$ as reduced matrices, i.e. after dividing each element by its (1,1) component. These plots show more clearly the behavior of the $\underline{R}_{aw}$ matrix elements at Brewster's angle. These curves agree exactly with the corresponding plots in Kattawar and Adams (1989, their Fig. 4).

Figure (3.4) shows $\underline{R}_{wa}$ and $\underline{T}_{wa}$. The vertical dotted line is at the critical angle for total internal reflection, which in the present case is $\theta_{\rm tir} = 48.36\,{\rm deg}$. For angles less than the critical angle, the transmission is never more than about 0.54. This again shows the n-squared law for radiance. In going from water to air, the in-water radiance is decreased by a factor of $1/n^2_{\rm water}$ when crossing the surface because the solid angle in air is greater than that in water by a factor of $n^2_{\rm water}$. The (1,1) elements show that beyond the critical angle there is no transmission and total reflection. These curves agree with the corresponding plots in Garcia (2012) (his Figs. 4-6).

Figure (3.5) shows the reduced water-to-air matrices. These curves agree with the corresponding plots in Kattawar and Adams (1989) (their Fig. 5)[1].

The non-zero matrix elements of course depend on incident angle as seen above, but also depend weakly on the wavelength via the wavelength dependence of $n_{\rm water}$.

The $\underline{R}_{aw}$ etc. matrices are non-dimensional reflectances and transmittances. The continuous-variable $\underline{r}_{aw}$ etc. seen in Eqs. 1.40 and 1.41 are radiance reflectances and transmittances with units of ${\rm sr}^{-1}$. We can add Dirac delta functions to the $\underline{R}_{aw}$ etc. to indicate the the exact directions of the reflected and transmitted light, which also gives ${\rm sr}^{-1}$ dimensions to the result. Let $\boldsymbol{\xi}(\theta,\phi)$ denote a unit vector pointing in direction (θ,ϕ). Then the delta function that "picks out" a particular direction $\boldsymbol{\xi}_o(\theta_o,\phi_o)$ is defined by

$$\delta(\boldsymbol{\xi}-\boldsymbol{\xi}_o) \equiv 0 \quad \text{if} \quad \boldsymbol{\xi} \neq \boldsymbol{\xi}_{\rm o}, \tag{3.33}$$

and

$$\int_{4\pi} f(\boldsymbol{\xi})\delta(\boldsymbol{\xi}-\boldsymbol{\xi}_o)d\Omega(\boldsymbol{\xi}) \equiv f(\boldsymbol{\xi}_o). \tag{3.34}$$

where $f(\boldsymbol{\xi})$ is any function of direction. Note that because the element of solid angle $d\Omega(\boldsymbol{\xi})$ has units of steradian, if follows that $\delta(\boldsymbol{\xi}-\boldsymbol{\xi}_o)$ has units of 1/steradian. When performing integrations over directions (μ,ϕ), $d\Omega(\boldsymbol{\xi}) = d\mu\, d\phi$ and the delta function becomes $\delta(\boldsymbol{\xi}-\boldsymbol{\xi}_o) = \delta(\mu-\mu_o)\,\delta(\phi-\phi_o)$.

We can thus write the surface transfer functions for a level sea surface as

$$\underline{r}_{aw} = \underline{R}_{aw}\,\delta(\boldsymbol{\xi}-\boldsymbol{\xi}_{r_{aw}}) \tag{3.35}$$

$$\underline{t}_{aw} = \underline{T}_{aw}\,\delta(\boldsymbol{\xi}-\boldsymbol{\xi}_{t_{aw}}) \tag{3.36}$$

$$\underline{r}_{wa} = \underline{R}_{wa}\,\delta(\boldsymbol{\xi}-\boldsymbol{\xi}_{r_{wa}}) \tag{3.37}$$

$$\underline{t}_{wa} = \underline{T}_{aw}\,\delta(\boldsymbol{\xi}-\boldsymbol{\xi}_{t_{wa}}) \tag{3.38}$$

Here $\boldsymbol{\xi}_{r_{aw}}$ is the direction of the air-incident light as reflected upward by the surface, $\boldsymbol{\xi}_{t_{aw}}$ is the direction of the light transmitted through the surface into the water, etc. The reflected directions are given by the law of reflection, $\theta = \theta'$. The transmitted directions are given by Snel's law, 3.1. For a level surface, the incident and final meridian planes and the reflection-transmission plane are all the same. Thus the rotation matrices reduce to identity matrices and are not shown.

[1]The signs of the $\underline{T}_{wa}(3,4)$ and $\underline{T}_{wa}(4,3)$ elements are reversed in the original Fig. 5, which according to Kattawar had a sign error.

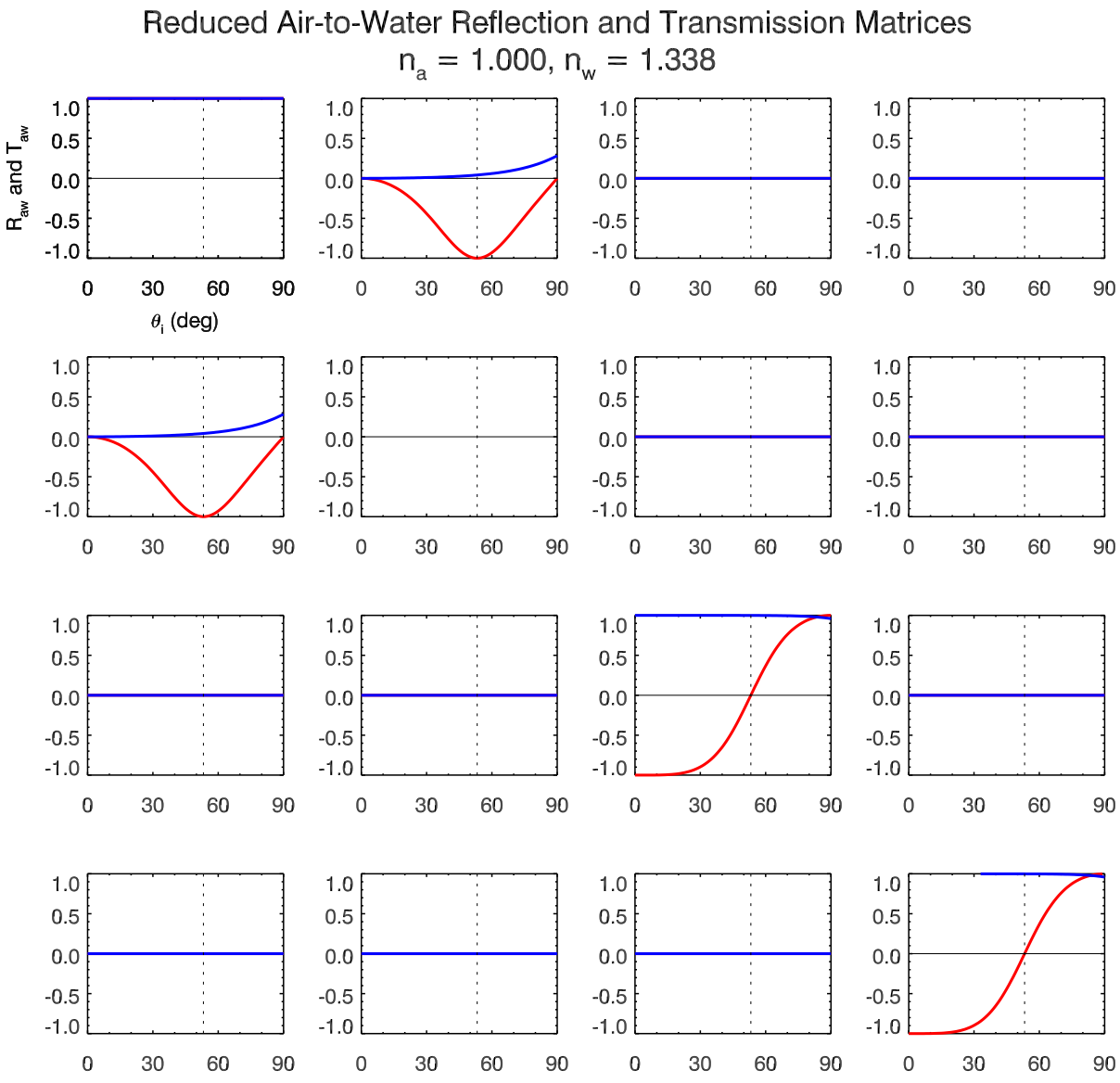

Figure 3.3: The reflectance and transmittance matrices of Fig.(3.2) normalized by their (1,1) elements. The vertical dotted line is Brewster's angle.

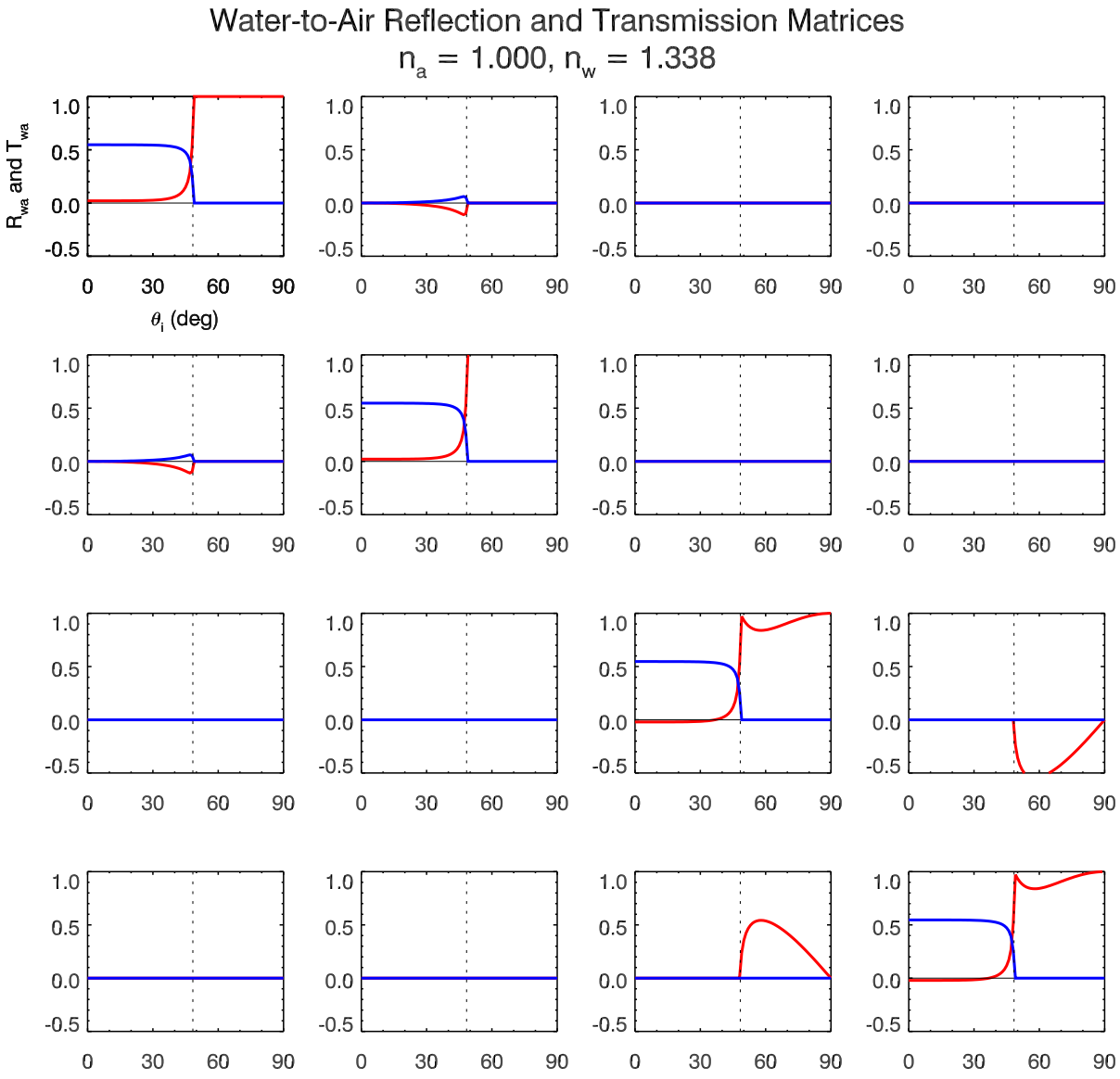

Figure 3.4: Reflectance and transmittance matrices as functions of the incident angle θ_i for water-incident radiance. $\underline{R}_{wa}$ is in red and $\underline{T}_{wa}$ is in blue. The vertical dotted line is the critical angle for total internal reflectance.

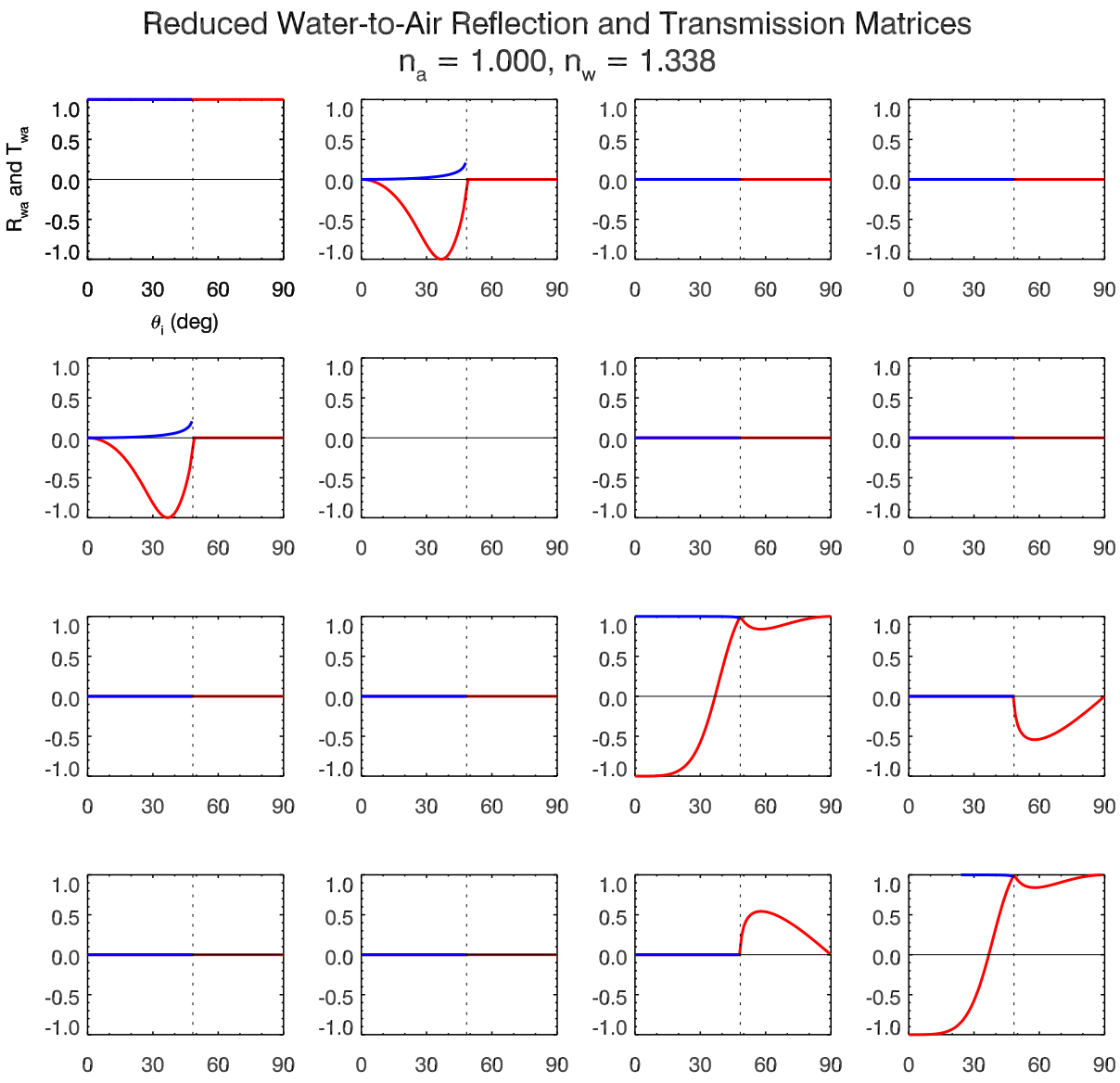

Figure 3.5: The reflectance and transmittance matrices of Fig.(3.4) normalized by their (1,1) elements. The 34 and 43 elements are the reverse of Fig. 5 in Kattawar and Adams (1989), which had a sign error in these two elements.

3.2 Reflection and Transmission by a Wind-blown Sea Surface

Wind-blown sea surfaces are vastly more complicated than level surfaces. The properties of the sea surface that determine its optical properties are the wave height and wave slope. The slope is of primary importance because the slope, along with the direction of an incident light ray, determines the angles of incidence and transmission seen in the Fresnel formulas. The wave height and slope in turn depend not just on the wind speed, but also on how long the wind has been blowing (the duration) and over what distance upwind from the point of interest (the fetch), on the direction of wave propagation relative to any current, on the possible presence of surface contaminants, on the depth in shallow water, and to a lesser extent even on air-sea temperature difference. Fortunately these details can be left to the physical oceanographers. This section shows how to compute the optical properties of a given sea surface, without regard for the environmental conditions that created that surface.

Recall that what HydroPol needs for its surface boundary conditions are the four discretized transfer functions $\underline{r}_{aw}$,..., $\underline{t}_{aw}$ seen in Eqs. (2.21) and (2.22). These transfer functions give the time- or space-averaged optical properties of wind-blown sea surfaces for a given set of environmental conditions. In practice, we can build up a library of these properties as functions of wind speed and water index of refraction. Then a given HydroPol run can select the appropriate set of surface transfer functions for the chosen wind speed and wavelength (the index of refraction depends on wavelength).

The surface transfer functions will be estimated by Monte Carlo ray tracing applied to numerous random realizations of the sea surface corresponding to a fixed set of environmental conditions such as wind speed. A random realization of the surface can be thought of as a snapshot of a patch of sea surface of some size, typically on the order of 10-100 m on a side. The surface elevations within this patch are known at a finely spaced grid of points, with perhaps 1000 or more points in each spatial direction. Ray tracing is used to compute the reflection and transmission properties of each realization for the full range of incident and final directions. Finally, averages over many such realizations–perhaps 10^5 or more–give the needed transfer functions.

We first show how sea surfaces can be modeled as a network of triangular wave facets, each of which is locally flat but randomly tilted from the horizontal. After generating a surface realization, we consider in detail how to trace rays through 3D space as they are reflected and refracted by the surface. This is a rather complicated business for two reasons. First, there can be multiple interactions of a ray with the surface because a ray reflected or refracted by one wave may intersect a neighboring wave some distance away, generating still more reflected and refracted rays. Second, each tilted wave facet has its own scattering plane for incident and scattered light, so careful attention must be paid to the rotations that carry the Stokes vector from one facet to another. We finally show how the Monte Carlo results lead to the needed transfer functions.

3.2.1 Modeling the Sea Surface as a Grid of Triangular Wave Facets

Figure 3.6 shows an idealized view of part of a water surface and a multiply scattered ray. A finite region of the mean water surface is resolved by a hexagonal grid of triangles, called *triads*. The sea surface elevation is defined at each triad vertex, so that the waves are represented by a set of triangular *wave facets*, which can be above or below mean sea surface. These facets are contained

in a hexagonal domain of 3D space defined by the boundaries of the hexagonal grid. Four such wave facets are shown in Fig. 3.6.

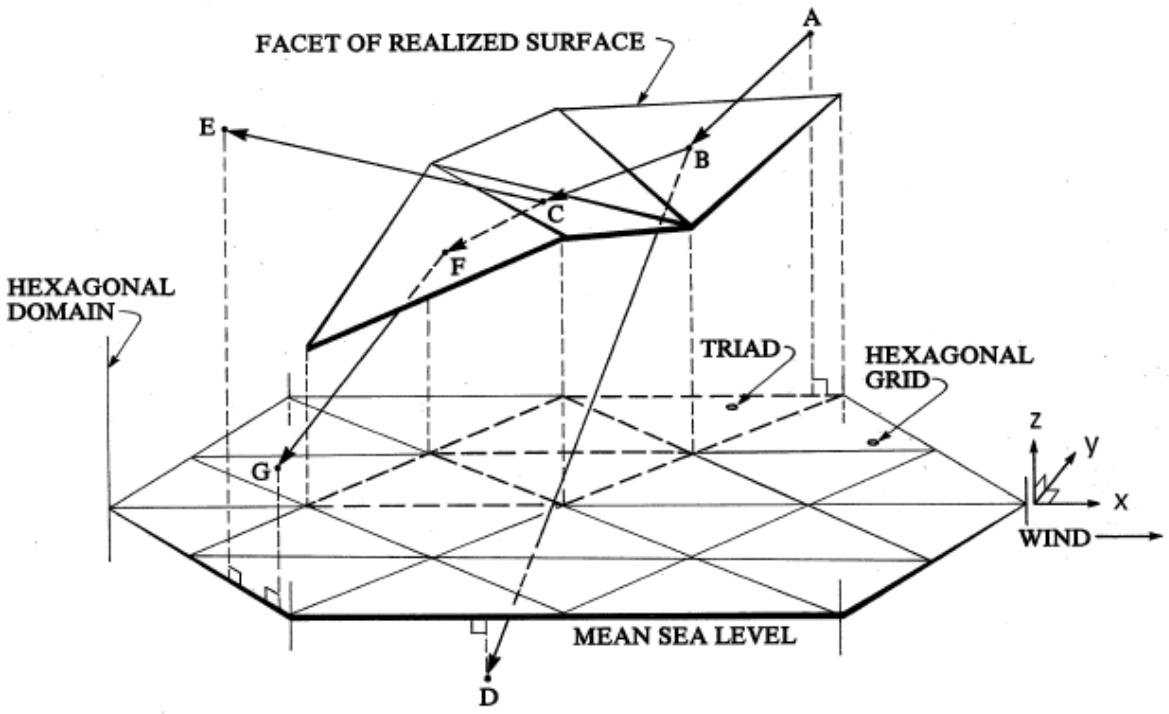

Figure 3.6: Illustration of the hexagonal domain of triangular wave facets for a particular surface realization and ray paths. Observe the orientation of the horizontal triads relative to the downwind direction.

The virtues of resolving the sea surface as a grid of triangles should be obvious. The three points of surface elevation for a wave facet determine a plane in 3D space, and the Fresnel equations of the previous section can be applied to that plane. If the surface were described by a rectangular grid of points, for example, then some additional decision would have to be made regarding the orientation in space of the point of intersection within the surface region specified by the four points giving the elevations of the rectangular wave element.

The hexagonal grid of Fig. 3.6 has $N_{\text{hex}} = 2$ "rings" or "layers" of triads around the center point of the grid. For actual simulations, many more triads are used, as will be shown below. A hexagonal grid with N_{hex} layers of triads has $N_{\text{nodes}} = 3N_{\text{hex}}(N_{\text{hex}} + 1)$ triad vertices where the surface elevations are defined.

There are several ways to generate surface elevations at the triad vertices, and we next describe the two most important. The simplest and most widely used method is based on the wind-speed wave-slope model of Cox and Munk (1954). The second is based on wave energy spectra.

Method 1: Surface Generation Based on Cox-Munk Statistics

The classic study by Cox and Munk (1954) used photos of sea-surface sun glint taken from aircraft to deduce the slope statistics of the surface waves for various wind speeds. The result is a statistical model of the alongwind and crosswind wave slopes as a function of wind speed. Their model describes the slopes due to whatever gravity and capillary waves were on the sea surface at the time the photos were taken. The virtue of the Cox-Munk wind-speed wave-slope equations is that they are very simple, and consequently they are widely used for sea surface generation (e.g., Zhai et al. (2010) and in HydroLight). The disadvantage of using their model is that the generated surface reproduces only the surface *slope* statistics, and not also the surface *height* statistics. The resulting surface is thus a rough surface with randomly sloping wave facets, but with the facet elevations never deviating much from the mean sea surface. Consequently, this model misses some wave features such as shadowing by large gravity waves, which can be important at near horizontal incident or viewing directions. Indeed, the Cox-Munk model breaks down for horizontal viewing directions (Zeisse, 1995). However, because proper modeling of wave slopes is the critical factor for computing the surface optical properties, the Cox-Munk model is often adequate for solar and viewing directions that are not near the horizon.

Let $\zeta(x, y)$ be the sea surface elevation at location (x, y). The wind is blowing in the $+\mathbf{x}$ direction. The alongwind and crosswind wave slopes are

$$\zeta_a = \frac{\partial \zeta}{\partial x} \quad \text{and} \quad \zeta_c = \frac{\partial \zeta}{\partial y}\,,$$

respectively. The Cox and Munk (1954) model says that these slopes are normally distributed with zero mean and alongwind and crosswind variances given by

$$\sigma_a^2 = s_a U \quad \text{where} \quad s_a = 3.16 \times 10^{-3} \text{ s m}^{-1} \tag{3.39a}$$

$$\sigma_c^2 = s_c U \quad \text{where} \quad s_c = 1.92 \times 10^{-3} \text{ s m}^{-1} \tag{3.39b}$$

where U is the wind speed at 12.5 m above mean sea level.

To construct a surface realization based on the Cox-Munk equations, we first pick a value of N_{hex}, which defines the number of triads in the hexagonal grid. Let the first triad, illustrated in Fig. 3.7, have vertices $\mathbf{v}_1, \mathbf{v}_2$ and $\mathbf{v}_3$, where the surface elevations are $\zeta_i = \zeta(x_i, y_i), i = 1, 2, 3$. Then the alongwind and upwind slopes for this triad are

$$\zeta_a = \frac{\Delta\zeta}{\Delta x} = \frac{\zeta_2 - \zeta_1}{\delta} \tag{3.40a}$$

$$\zeta_c = \frac{\Delta\zeta}{\Delta y} = \frac{\zeta_3 - \frac{1}{2}(\zeta_1 + \zeta_2)}{\epsilon} \tag{3.40b}$$

where δ and ϵ are defined in Fig. 3.7. Let $\langle \varkappa \rangle$ denote the ensemble average of $\varkappa$, i.e. the average of $\varkappa$ over all of the surface realizations in the ensemble. Then by construction

$$\langle \zeta_j \rangle = 0 \quad \text{and} \quad \langle \zeta_i \zeta_j \rangle = \sigma^2 \delta_{i-j}\,,$$

where $i, j = 1, 2, 3$ and δ_k is the Kronecker delta symbol defined in Eq. (A.1). Similarly, it follows that ζ_a and ζ_c are distributed with zero means and variances

$$\sigma_a^2 = \langle \zeta_a^2 \rangle = \frac{2\sigma^2}{\delta^2} \quad \text{and} \quad \sigma_c^2 = \langle \zeta_c^2 \rangle = \frac{3\sigma^2}{2\epsilon^2}\,, \tag{3.41}$$

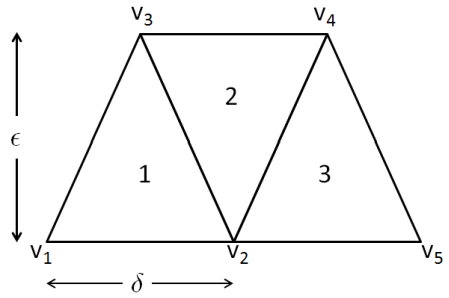

Figure 3.7: Triad dimensions and vertices as used for surface generation.

and that ζ_a and ζ_c are uncorrelated: $\langle \zeta_a \zeta_c \rangle = 0$. The elevation of the centroid of the facet is

$$\zeta = \frac{\zeta_1 + \zeta_2 + \zeta_3}{3} .$$

This ζ is normally distributed with zero mean and variance

$$\sigma_\zeta^2 = \langle \zeta^2 \rangle = \frac{\sigma^2}{3} . \tag{3.42}$$

We can now fix the triad sizes δ and ϵ, and the vertical scale σ, as a function of the wind speed U. From Eqs. (3.39) and (3.41) we have

$$\frac{\epsilon^2}{\delta^2} = \frac{3 s_a}{4 s_c} . \tag{3.43}$$

This means that the *shape* of the triad is independent of the wind speed and the physical units of δ and ϵ. That is to say, the slopes of the facets are independent of their physical size. The value of σ is obtained from Eqs. (3.39) and (3.41),

$$\sigma^2 = \frac{1}{2} s_a \delta^2 U . \tag{3.44}$$

Thus the vertical scale of the waves depends on the wind speed (the waves become more choppy as the wind increases).

Given the independence of the slopes on the triad physical dimensions, we are free to set $\delta = 1$ and generate the surface wave elevations using

$$\delta = 1 \tag{3.45a}$$

$$\epsilon = \left(\frac{3 s_a}{4 s_c} \right)^{1/2} \tag{3.45b}$$

$$\sigma = \left(\frac{s_a}{2} U \right)^{1/2} \tag{3.45c}$$

where the values of s_a and s_c are given in Eq. (3.39)

To generate a surface grid based on the Cox-Munk slope model, we first use Eqs. (3.45a) and (3.45b) to lay out the (x, y) locations of the triad vertices. Then to define the surface elevations for the first triad, labeled 1 in Fig. in Fig. 3.7, we draw three random numbers from a normal distribution with 0 mean and variance σ^2, denoted $\mathcal{N}(0, \sigma^2)$, with σ given by the wind speed U and Eq. (3.45c). These random numbers are the surface elevations ζ_1, ζ_2, and ζ_3 at triad vertices

$\mathbf{v}_1$, $\mathbf{v}_2$, and $\mathbf{v}_3$. To define the next facet, number 2 in the figure, only one new random number is drawn, which gives the elevation ζ_4 at vertex $\mathbf{v}_4$. The previously computed elevations at $\mathbf{v}_2$, and $\mathbf{v}_3$, and this ζ_4 at $\mathbf{v}_4$ then define facet 2. Each subsequent facet requires only one new random number, e.g., $\mathbf{v}_2$, and $\mathbf{v}_4$ plus a new ζ_5 at $\mathbf{v}_5$ define facet 3, and so on. Sea surface wave facets defined in this manner then obey the slope statistics given by the Cox-Munk model of Eqs. (3.39).

Figure 3.8 is an example of a sea surface generated as a hexagonal grid of triangluar wave facets using the Cox-Munk slope statistics for a wind speed of $U = 6$ m s^{-1}. The x, y, z values are scaled in terms of the triad x-dimension $\delta = 1$. The notable feature of this surface is that all waves have a horizontal scale of order a few δ in size. That is, there is no obvious superposition of both large- and small-scale waves, as is typically the case of a real ocean surface. With this method, the wave elevations are simply whatever they need to be so that the facets reproduce the Cox-Munk wave slope statistics.

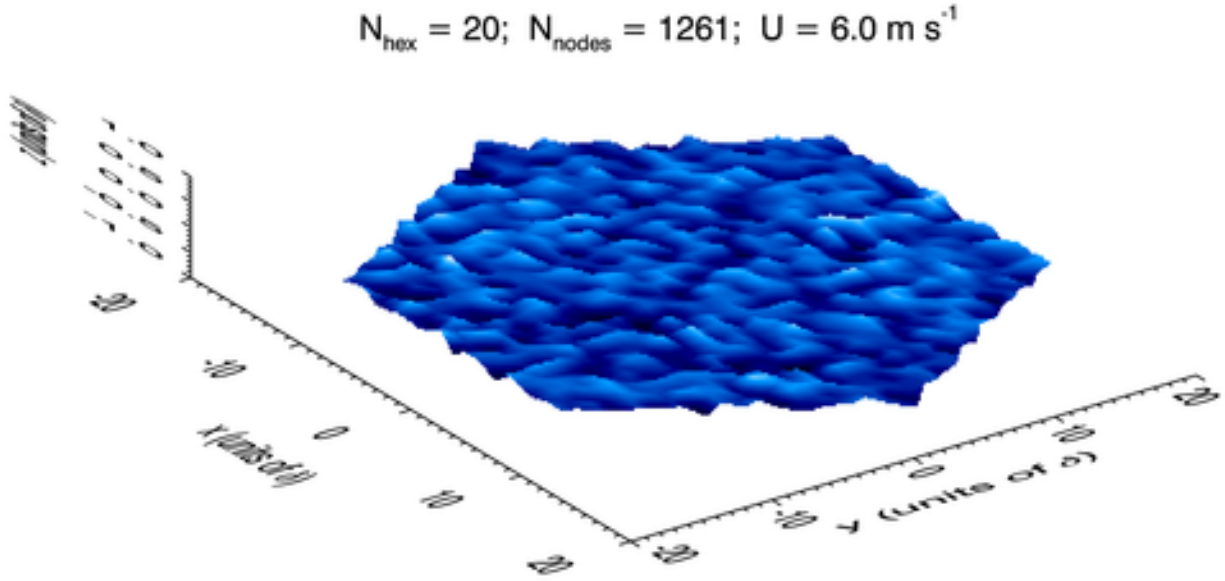

Figure 3.8: Example of a sea surface generated using the Cox-Munk wind-speed wave-slope statistics and a hexagonal grid of triangular wave facets. Light blue indicates wave crests (large positive values) and dark blue is wave troughs (large negative values). The graphics program used to create this visual rendering of the surface causes some smoothing, so the individual wave facets are not apparent.

Method 2: Surface Generation Based on Wave Energy Spectra

More realistic sea surface realizations that model the wave height, from which the slopes are obtained directly, can be obtained at the expense of doing more physics and mathematics. Without going into the details, the foundation of this process is as follows. Suppose that we have measured the sea surface elevation $\zeta(x_r, y_s) \equiv \zeta(r, s)$ at a given time and over a rectangular grid of spatial points $(x_r, y_s), r = 1, ..., N_x, s = 1, ..., N_y$. This grid covers an area of some physical size $X \times Y$, e.g. $0 \le x \le X, 0 \le y \le Y$, with a grid spacing of $\Delta x = X/N_x$ and $\Delta y = Y/N_y$. The longest

wavelengths resolvable by this grid are X and Y in the respective directions, and the smallest waves are of wavelength $2\Delta x = 2X/N_x$ and $2\Delta y = 2Y/N_y$. The surface comprises waves of all wavelengths between these two limits.

Linear wave theory says that this surface can be represented as a sum of sinusoids:

$$\zeta(x_r, y_s) = \frac{a_0}{2} + \sum_{u=1}^{N_x/2} \sum_{v=1}^{N_y/2} \left[a_{u,v} \cos(k_u x_r + k_v y_s) + b_{u,v} \sin(k_u x_r + k_v y_s) \right] , \tag{3.46}$$

where the $a_{u,v}$ and $b_{u,v}$ are the amplitudes of the waves with spatial frequencies $k_u = 2\pi u/X$ and $k_v = 2\pi v/Y$; k_u and k_v have units of radian per meter (a_0 represents mean sea level and can be set to 0). This decomposition can be placed in the form of a discrete Fourier transform pair:

$$\zeta(r, s) = \sum_{u=0}^{N_x-1} \sum_{v=0}^{N_y-1} \hat{\zeta}(u, v) e^{+i2\pi(ru/N_x + sv/N_y)} \tag{3.47}$$

$$\hat{\zeta}(u, v) = \frac{1}{N_x N_y} \sum_{r=0}^{N_x-1} \sum_{v=0}^{N_y-1} \zeta(r, s) e^{-i2\pi(ru/N_x + sv/N_y)} . \tag{3.48}$$

Equation (3.48) is usually called the discrete Fourier transform, and Eq. (3.47) is the inverse transform. (These equations can be written in various ways. The form used here is that of Bracewell (1986).) Note that although $\zeta(r, s)$ is real, $\hat{\zeta}(u, v)$ is in general a complex number. However, $\hat{\zeta}(u, v)$ is Hermitian, which means that $\hat{\zeta}^*(u, v) = \hat{\zeta}(-u, -v)$, where * denotes the complex conjugate. The Hermitian-ness of $\hat{\zeta}(u, v)$ guarantees that when used in Eq. (3.47), the resulting $\zeta(r, s)$ is real. These matters can be pursued in texts on Fourier transforms, e.g., Bracewell (1986).

The energy of a wave is proportional to the square of its amplitude. Thus the square of $\hat{\zeta}(u, v)$ is proportional to the energy contained in waves with frequencies of k_u, k_v in the respective x, y directions. The quantity $\mathcal{E}(k_u, k_v) = \|\hat{\zeta}(u, v)\|^2 = \hat{\zeta}(u, v)\hat{\zeta}^*(u, v)$ is therefore known as the wave energy spectrum. $\mathcal{E}(k_u, k_v)$ shows how the wave energy is partitioned into waves of different spatial frequencies, which is of great interest to physical oceanographers. Much work has therefore been done to develop models of wave energy spectra as functions of wind speed, fetch, and duration. These models can be used to generate random realizations of surface waves such that the generated surfaces are consistent with the chosen energy spectrum.

One way this can be done is given in §4.5 of Mobley (1994). However, Eq. (4.77) of that work amounts to a brute-force evaluation of Eq. (3.46) or (3.47) and is much too computationally intensive for use in generating the thousands of wave surfaces for large spatial regions and and large N_x, N_y. However, if N_x and N_y are both powers of 2, then the sums of Eqs. (3.47) and (3.48) can be evaluated using the Fast Fourier Transform (FFT) algorithm, and the computer time is orders of magnitude less.

To use a given energy spectrum $\mathcal{E}(k_u, k_v)$ to generate a random surface realization $\zeta(r, s)$, we first choose the X, Y and N_x, N_y values, which determine the physical size of the spatial region and the fineness of the wave resolution. This gives an x-y grid of points at which $\zeta(r, s)$ is to be determined. Next choose an energy spectrum $\mathcal{E}(k_u, k_v)$, which will be a complicated set of equations depending on wind speed and perhaps other environmental parameters such as fetch. Then for each (u, v) combination, draw independent random numbers $\rho_1(u, v)$ and $\rho_2(u, v)$ from a

$\mathcal{N}(0,1)$ distribution, and form the quantity

$$\hat{\zeta}_o(u,v) = \frac{1}{\sqrt{2}}[\rho_1(u,v) + i\rho_2(u,v)]\sqrt{\frac{\mathcal{E}(k_u,k_v)}{2}}\,. \tag{3.49}$$

Here $i = \sqrt{-1}$ and all other quantities are real. The expectation of $\|\hat{\zeta}_o\|^2$ is (omitting the (u,v) argument for brevity)

$$\begin{aligned}\langle\|\hat{\zeta}_o\|^2\rangle = \langle\hat{\zeta}_o\hat{\zeta}_o^*\rangle &= \langle\left[\frac{1}{\sqrt{2}}(\rho_1 + i\rho_2)\sqrt{\mathcal{E}/2}\right]\left[\frac{1}{\sqrt{2}}(\rho_1 - i\rho_2)\sqrt{\mathcal{E}/2}\right]\rangle \\ &= \frac{\mathcal{E}}{2}\left(\langle\rho_1^2\rangle + \langle\rho_2^2\rangle\right) = \mathcal{E}\end{aligned}$$

because $\langle\rho_1^2\rangle = \langle\rho_2^2\rangle = 1$ for $\mathcal{N}(0,1)$ random variables, and $\langle\rho_1\rho_2\rangle = 0$ because these random variables are uncorrelated. Thus this $\hat{\zeta}_o$ is consistent with the chosen energy spectrum. However, $\hat{\zeta}_o(u,v)$ is not Hermitian, so using $\tilde{\zeta}_o(u,v)$ in Eq. (3.47) would not give a real sea surface.

However, the quantity

$$\hat{\zeta}(u,v) = \frac{1}{2}[\hat{\zeta}_o(u,v) + \hat{\zeta}_o^*(u,v)] \tag{3.50}$$

is Hermitian. An analysis like that for $\langle\|\hat{\zeta}_o\|^2\rangle$ shows that $\langle\|\hat{\zeta}\|^2\rangle = \mathcal{E}$. Thus the $\hat{\zeta}(u,v)$ of Eq. (3.50), when used in Eq. (3.47), gives a sea surface realization that is real and consistent with the chosen wave energy spectrum.

Figure 3.9 shows an example of a random surface generated in this manner. The spatial region is $X \times Y = 25 \times 25$ m, which is resolved by $N_x \times N_y = 2048 \times 1024$ grid points. The smallest resolved wavelength in the x direction is then $2 \times 25/2048 = 0.024$ m, and twice that in the crosswind direction. Thus this particular surface resolves all wavelengths from 25 m gravity waves almost down to capillary waves, which have wavelengths of order 1 cm. The energy spectrum was that of Elfouhaily et al. (1997) with a wind speed of 6 m s^{-1} and other parameters set to values for a fully developed sea. The computer time required for generation of this surface by the FFT evaluation of Eq. (3.47) was only 0.56 seconds even though there were almost 2.1×10^6 surface elevations generated. (This time is for the IDL FFT routine on a 2.4 GHz computer.) The visual appearance of this surface is much different than for the Cox-Munk surface of Fig. 3.8. In particular, it is now obvious that the surface is a superposition of waves of different horizontal scales, even though the smallest-scale waves are not visually apparent in this figure.

Although the sea surface of Fig. 3.9 was rapidly generated and resolves a wide-range of wavelengths, the surface elevations are defined on a rectangular x-y grid, not on the triangular grid used for ray tracing. However, this is not a problem. Figure 3.10 shows how a rectangular FFT grid can be mapped to a hexagonal grid of triangular wave facets. It may seem wasteful to compute the surface elevation at each of the blue grid points (the points where the blue lines intersect) and then discard roughly half of these points to subsample the FFT grid to obtain a hexagonal grid. However the extremely fast run time of the FFT algorithm makes this option much more efficient than the regular Fourier technique given in Mobley (1994) (his Eq. 4.77). That technique computes the surface elevations only at the triad vertices, but requires much more time to evaluate the equivalent of Eq. (3.46), in which the $a_{u,v}, b_{u,v}$ coefficients are randomly determined from the chosen wave energy spectrum. Figure 3.10 also makes clear why N_y was chosen be $N_x/2$ in the generation of the FFT grid. $N_y = N_x/2$ allows every row of y values to be used, even though

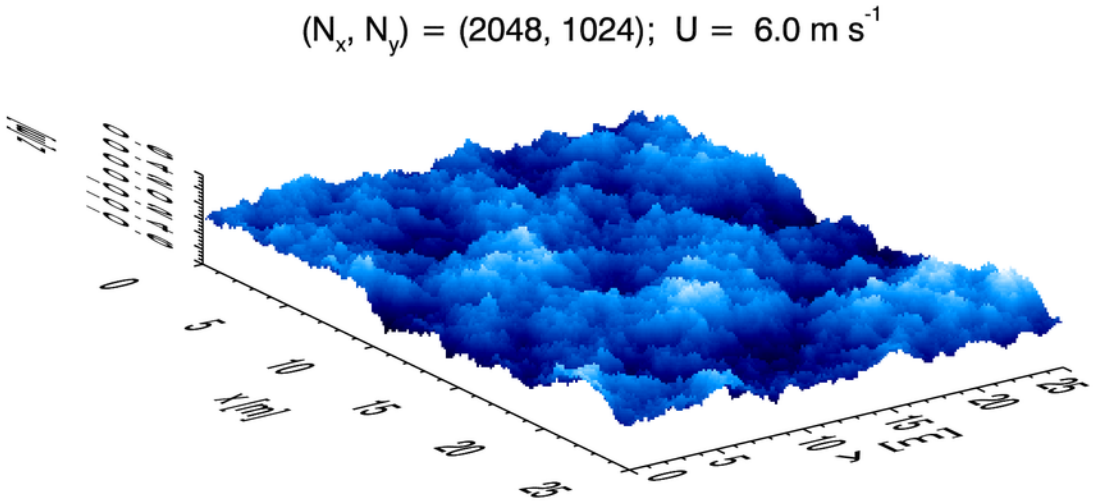

Figure 3.9: Example sea surface generated by FFT techniques. Light blue is high elevations and dark blue is low. The surface appears visually rougher than a real sea surface because of the expanded vertical scale relative to the horizontal dimensions.

only every other x value is retained when generating the hexagonal grid. Figure 3.11 shows the sea surface of Fig. 3.9 after subsampling to create a hexagonal grid of wave facets. Visually, the original and subsampled surfaces are indistinguishable at the scale of these figures.

Sea surface generation using FFT techniques is widely used, especially in creating computer-animated movies. However, the devil is in the details. When computing $\hat{\zeta}(u,v)$ from sea surface elevations via Eq. (3.48), $\hat{\zeta}(u,v)$ has both positive and negative spatial frequencies. These correspond to waves traveling both "left" and "right." In essence, with just a snapshot $\zeta(r,s)$ of the sea surface, the mathematics cannot tell which direction the waves are going. The resulting energy spectrum $\mathcal{E}(k_u,k_v)$ also has positive and negative k_u, k_v, with $\mathcal{E}(-k_u,-k_v) = \mathcal{E}(k_u,k_v)$, and is termed a "two-sided" spectrum. However, wave spectra are usually presented as "one-sided" spectra showing only the positive frequencies, but with double the magnitude of the two-sided $\mathcal{E}$. One-sided spectra show how much energy is contained in a give frequency interval, without regard to the sign of the k_u or k_v argument, i.e., without regard for the direction the waves are traveling relative to a given axis. Thus the one-sided $\mathcal{E}$ used in Eq. (3.49) is divided by 2. The widely cited tutorial by Tessendorf (2004) gives equations corresponding to Eqs. (3.49) and (3.50) (his Eqs. 41 and 43), but without dividing $\mathcal{E}$ by 2, and without the factor of $\frac{1}{2}$ in Eq. (3.50). In a round-trip calculation from energy spectrum to sea surface and back to energy spectrum, his equations do not conserve energy by a factor of $(2\sqrt{2})^2 = 8$. That is, his versions of Eqs. (3.49) and (3.50) generate wave amplitudes that are too large. This is acceptable for the movie industry, which has never worried about conserving energy or exceeding the speed of light, but it isn't good enough for science. There are other subtleties such as tricks of array storage regarding how FFT routines order the frequencies. These matters need not concern us here, but do require care in the computer

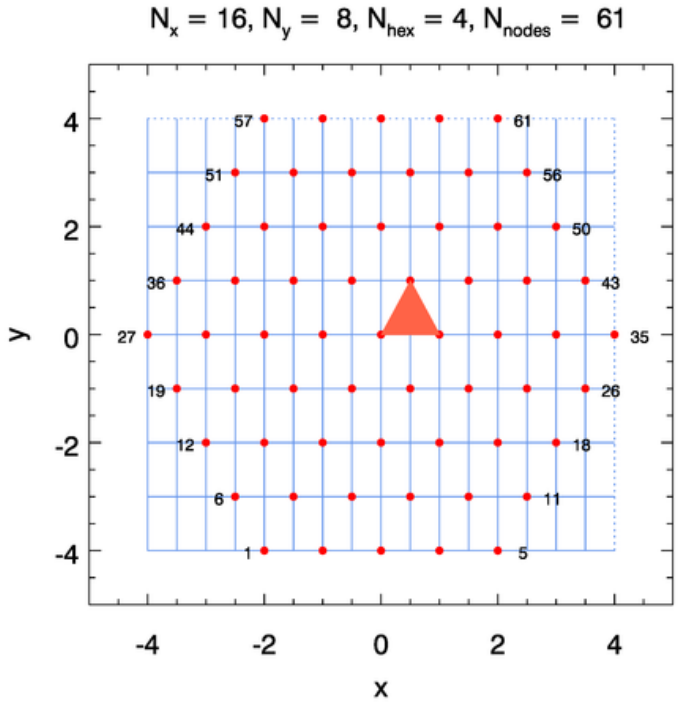

Figure 3.10: Example mapping of a rectangular computational grid used in FFT simulations of a random sea surface (blue lines) to a hexagonal grid of triangular wave facets used in ray tracing (red dots). This example has $N_x = 16$ points in the x direction and $N_y = 8$ in the y direction for the FFT calculations. The FFT grid points at the right and top, shown by the dotted blue lines, are obtained by the inherent spatial periodicity of the surface as determined by FFT techniques. Thus the surface elevation of point 35 is the same as that of point 27, point 57 is the same as point 1, etc. One of the triads generated from the FFT grid is shaded in red.

programming associated with surface generation using FFT techniques.

3.2.2 Ray Tracing in a Grid of Triangular Wave Facets

After a surface has been generated–i.e., after the elevations are defined at each triad vertex–a light ray described by a known Stokes vector is aimed toward the surface from any chosen direction. Figure 3.6 shows such a ray entering the hexagonal domain at point A. Every such initial ray eventually strikes a surface wave facet, as at point B. Each encounter of a ray with a wave facet usually generates both a reflected and a refracted daughter ray, the exception being total internal reflection, which generates only a reflected ray. The directions and radiant properties of these daughter rays are determined by Snel's law and the Fresnel equations applied to the plane surface of the facet. The incident angle onto the facet is computed from the incident ray direction and the direction of the normal to the wave facet. The daughter rays may undergo further encounters with other wave facets. As illustrated in Fig. 3.6, the first refracted ray at B, heading downward through the water, leaves the hexagonal domain at D without further scattering. The first reflected ray at

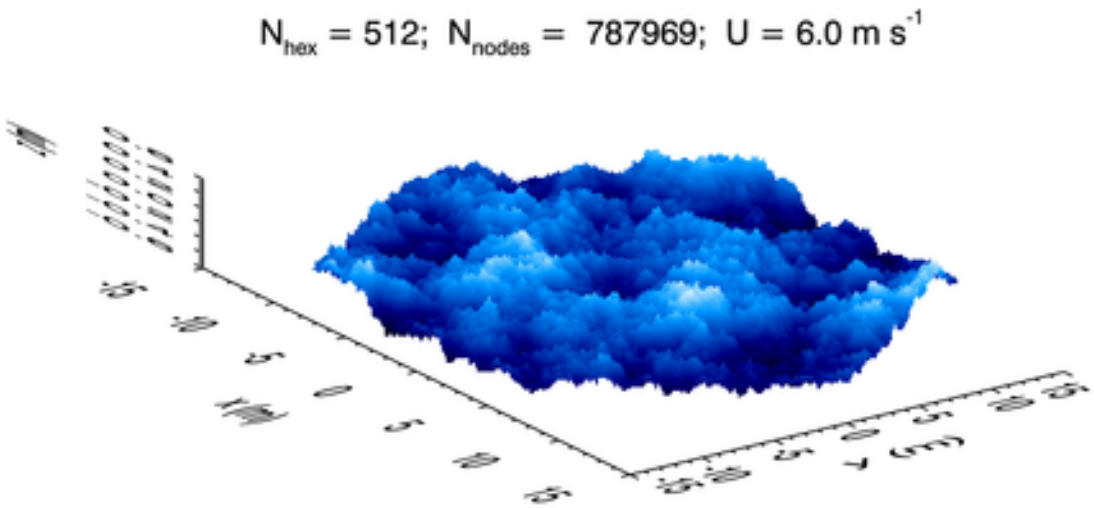

Figure 3.11: The rectangular FFT grid of Fig. 3.9 after conversion to a hexagonal grid and recentering to $(x, y) = (0, 0)$ at the origin of the hexagon.

B, however, intercepts another facet at C, generating two more rays. The reflected ray starting from C leaves the domain at E. The refracted ray starting from C encounters yet another facet at F and undergoes a total internal reflection before leavint the domain at G. Thus this initial ray finally results in one reflected and two refracted rays emerging from the hexagonal domain. The radiant properties at the points of emergence contribute to the corresponding reflectance and transmittance properties connecting the incident and final directions.

The interactions between random surface facets and rays incident from any direction lead to a proliferation of daughter rays that cannot be predicted in advance. These rays are processed as follows. Whenever a ray is created, either as an initial ray or as a daughter ray, the ray information is immediately "pushed" into, i.e. stored in, a two-dimensional array $\mathcal{S}$, called the *stack*. The information recorded in the stack includes

- the point $\mathbf{p}$ of the ray's origin,
- the ray's direction $\boldsymbol{\xi}$,
- a vector $\mathbf{s}$, which perpendicular to the meridian or scattering plane used to specify the Stokes vector at $\mathbf{p}$,
- a 4×4 matrix $\underline{\mathcal{W}}$. $\underline{\mathcal{W}}$ is essentially a weighting function that describes the ray's current radiant properties; $\underline{\mathcal{W}}$ is described in §3.2.2.

The dimensions of the stack array are $(N_{\text{maxrays}}, 25)$. N_{maxrays} is the maximum number of rays that can be stored at once. Experience with stack arrays in the HydroLight code shows there there are never more than 10 rays in the stack awaiting processing at any one time during ray tracing. Therefore we can set $N_{\text{maxrays}} = 10$. The dimension of 25 is determined by the 3 numbers for the

initial point, the 3 for the ray direction, 3 for the perpendicular vector direction, and 16 for the elements of the weight matrix.

During processing the next available ray in the stack is "pulled" from the stack and traced to completion. Completion occurs either when the ray leaves the hexagonal domain or when it encounters a wave facet. If the ray leaves the domain, its properties are tallied to the accumulating estimate. If the ray encounters a wave facet, the daughter ray or rays are "pushed" into the stack to away further processing. After completion of one ray, the next one is pulled from the stack and processed. This use of a push-pull stack allows the number of rays to grow as needed, according to the interactions of rays with the random surface. The stack will eventually be emptied, at which time a new initial ray is created. Figure 3.12 shows a flowchart of these operations.

Tracing the path of a light ray through 3D space as it is reflected and refracted by the sea surface is in principle no more complicated than determining where a straight line intersects a plane. Those calculations are purely geometric and do not depend on the polarization state of the light. However, the application of the basic geometric concepts to the irregular geometry of the wave facets comprising a random realization of the sea surface is rather tedious. The details of those geometric calculations are given in Appendix B.

When the light ray being traced represents polarized light, careful account must be taken of the rotations that carry the Stokes vectors from one scattering event to the next. The Stokes vector $\underline{S}_i$ of an incident ray is specified in the incident meridian plane, defined by the $\mathbf{z}$ axis and the direction $\boldsymbol{\xi}'$ of the ray. When that ray intersects a surface wave facet, the incident direction and the local normal $\mathbf{n}$ to the surface facet determine the plane containing the incident, reflected, and refracted directions. This is a local scattering plane in the sense of Fig. 1.3. The incident Stokes vector must be rotated from the incident meridian plane to the scattering plane of the facet. The Fresnel equations of the preceding §3.1 can then be applied to compute the polarization states of the reflected and refracted rays. If those rays leave the hexagonal domain, they are then rotated into the final meridian plane. In the case of a single scattering by a wave facet, the rotation angles can be determined from the spherical triangle defined by the incident and final meridian planes and the plane of reflection and refraction. This is the same geometry as the scattering situation shown in Fig. 1.3.

However, for scattering by a wind-blown sea surface, many rays will intersect another wave facet, perhaps several times, before leaving the hexagonal domain. Thus rays may go from one wave facet to another without being referred to any meridian plane (that could be done, but would cause unneeded calculations). We thus need an algorithm to determine the scattering planes of intersected facets, along with the associated rotation matrices that carry the Stokes vectors from one wave facet to another. Only the rays leaving the hexagonal domain need to referred to a meridian plane. This section describes the vector analysis needed to determine these facet-to-facet scattering planes and rotations. We assume here that the ray paths and points of facet intersection are known from the calculations described in Appendix B.

Figure 3.13 illustrates a single scattering of an air-incident ray by a surface wave facet. Air-incident means that the ray intersects the surface facet from the air side. Such a ray is always traveling downward for an initial ray, but may also be traveling upward for rays that have been scattered by a previous wave facet (as will be seen in Fig. 3.15 below). At the start of ray tracing, we know the incident ray direction $\boldsymbol{\xi}_i$. This direction, along with the $\mathbf{z}$ axis, defines the incident

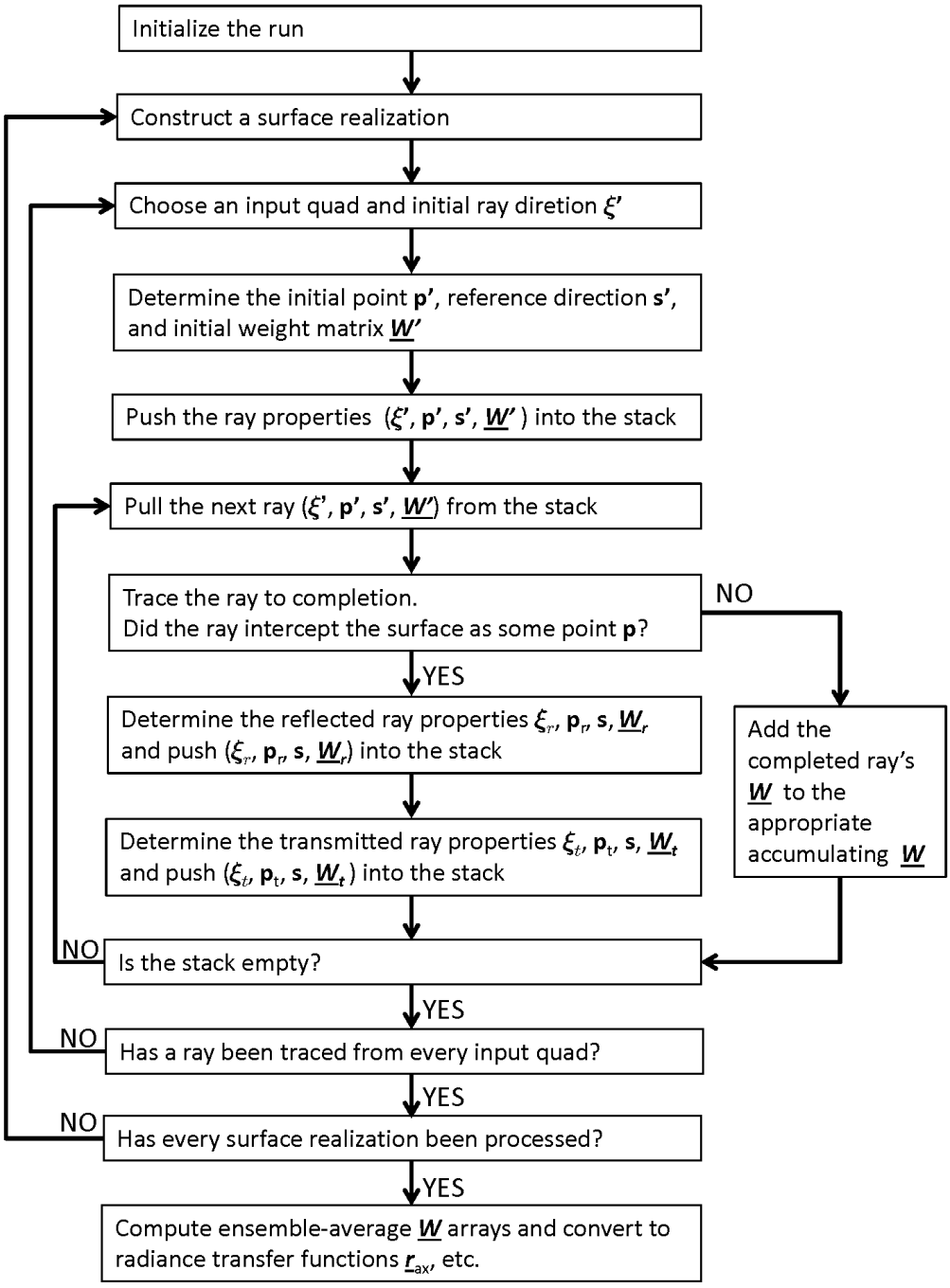

Figure 3.12: Flow chart of the Monte Carlo ray tracing computation of the sea surface transfer functions.

meridian plane, which is shown by a red dotted line in Fig. 3.13 connecting the pole, the tail of the $\boldsymbol{\xi}_i$ vector as drawn, and the equator. The associated Stokes vector $\underline{S}_i$ is specified relative to horizontal and vertical directions, $\mathbf{h}_i$ and $\mathbf{v}_i$, which are respectively perpendicular and parallel to the incident meridian plane. Vectors $\mathbf{h}_i$ and $\mathbf{v}_i$ are determined so that "perpendicular" cross "parallel" is in the direction of light propagation. For an incident ray this rule gives

$$\mathbf{h}_i = \frac{\mathbf{z} \times \boldsymbol{\xi}_i}{|\mathbf{z} \times \boldsymbol{\xi}_i|}\,, \tag{3.51a}$$

$$\mathbf{v}_i = \boldsymbol{\xi}_i \times \mathbf{h}_i\,, \tag{3.51b}$$

$$\boldsymbol{\xi}_i = \mathbf{h}_i \times \mathbf{v}_i\,. \tag{3.51c}$$

(For the degenerate case of $\boldsymbol{\xi}_i = \pm\mathbf{z}$, we are free to set $\mathbf{h}_i = \mathbf{x}$ and $\mathbf{v}_i = \pm\mathbf{y}$.) Vectors $\mathbf{h}_i$ and $\mathbf{v}_i$ and shown by green arrows in the figure. The green dotted line passing through the tail of the $\boldsymbol{\xi}_i$ vector makes it easy to see that this $\mathbf{h}_i$ is in the $-\boldsymbol{\varphi}$ direction as given by Eq. (1.2b). Vector $\mathbf{v}_i$ is parallel to the meridian plane, and is in the $-\boldsymbol{\vartheta}$ direction.

As was explained in the discussion of rotation angles in §1.2.2, we could just as well define $\mathbf{h}_i = \boldsymbol{\xi}_i \times \mathbf{z}/|\boldsymbol{\xi}_i \times \mathbf{z}|$. This would reverse the directions of $\mathbf{h}_i$ and $\mathbf{v}_i$ in Fig. 3.13, but that gives the same Stokes vector components because rotations by 180 deg leave Stokes vectors unchanged.

The incident vector always intersects a wave facet, which is illustrated by the blue triangle in Fig. 3.13. The unit outward normal to the wave facet is $\mathbf{n}$. The vertices of the wave facet are known from the ray tracing algorithm of Appendix B, from which $\mathbf{n}$ is given by Eq. (B.16). The facet normal $\mathbf{n}$ is tilted from the normal to the mean sea surface, $\mathbf{z}$, by polar and azimuthal angles (θ_n, ϕ_n). These angles are related to $\mathbf{n}$ by Eqs. (1.3a)-(1.3c):

$$\begin{aligned} \theta_n &= \cos^{-1} n_z \\ \phi_n &= \tan^{-1}(n_y/n_x)\,. \end{aligned}$$

The incident direction $\boldsymbol{\xi}_i$ and the facet normal $\mathbf{n}$ define the scattering plane containing the incident, reflected ($\boldsymbol{\xi}_r$) and transmitted ($\boldsymbol{\xi}_t$) rays. The plane containing these four rays is indicated by the dashed circle in Fig. 3.13, and the part of that plane containing $\boldsymbol{\xi}_i$ as drawn, $\mathbf{n}$, and $\boldsymbol{\xi}_r$ is shaded in red. The incident Stokes vector must be rotated from the incident meridian plane into a coordinate system whose axes that are normal ($\mathbf{s}$) and parallel ($\mathbf{p}$) to the scattering plane. Using the choices of incident direction cross facet normal equals the perpendicular vector, and perpendicular cross parallel equals direction of propagation, give

$$\mathbf{s} = \frac{\boldsymbol{\xi}_i \times \mathbf{n}}{|\boldsymbol{\xi}_i \times \mathbf{n}|}\,, \tag{3.52a}$$

$$\mathbf{p} = \boldsymbol{\xi}_i \times \mathbf{s}\,, \tag{3.52b}$$

$$\boldsymbol{\xi}_i = \mathbf{s} \times \mathbf{p}\,. \tag{3.52c}$$

Vector $\mathbf{s}$ lies in the plane of the wave facet and is is normal to each of $\boldsymbol{\xi}_i$, $\mathbf{n}$, $\boldsymbol{\xi}_r$, and $\boldsymbol{\xi}_t$. Vector $\mathbf{p}$ lies in the scattering plane and is normal only to $\boldsymbol{\xi}_i$.

The rotation angle $0 \le \alpha_i < 2\pi$ that rotates the incident Stokes vector into the scattering plane is the angle that rotates the initial perpendicular direction $\mathbf{h}_i$ into the direction perpendicular to the scattering plane. (This same angle rotates $\mathbf{v}_i$ into $\mathbf{p}$.) The rotation angle can therefore be obtained

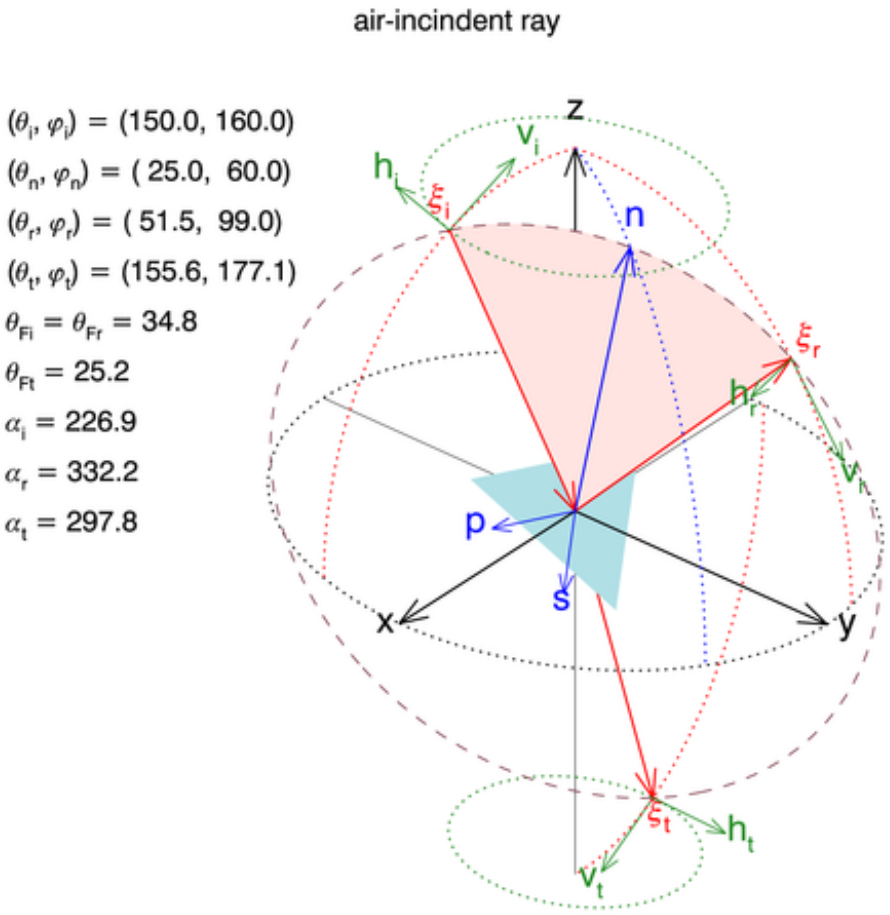

Figure 3.13: Directions and angles for an air-incident ray. Red arrows show the incident, reflected, and transmitted ray directions. The respective meridian planes are shown by red dotted lines. The light blue triangle represents the wave facet, whose normal **n** is shown in blue. The three sets of green arrows show the horizontal and vertical directions for specification of the Stokes vectors in the incident and final meridian planes. The maroon dashed line connecting the tail of $\boldsymbol{\xi}_i$ and the heads of **n**, $\boldsymbol{\xi}_r$, and $\boldsymbol{\xi}_t$ identifies the scattering plane containing the incident, reflected, and refracted directions; part of the scattering plane is shaded in red. Angles (θ, ϕ) give the directions of vectors with respect to the **x**-**y**-**z** axes. The $\theta_{\rm F}$ angles give the incident, reflected, and transmitted directions relative to the facet normal **n**; these are the angles used in the Fresnel formulas of §3.1. The α angles are the rotation angles.

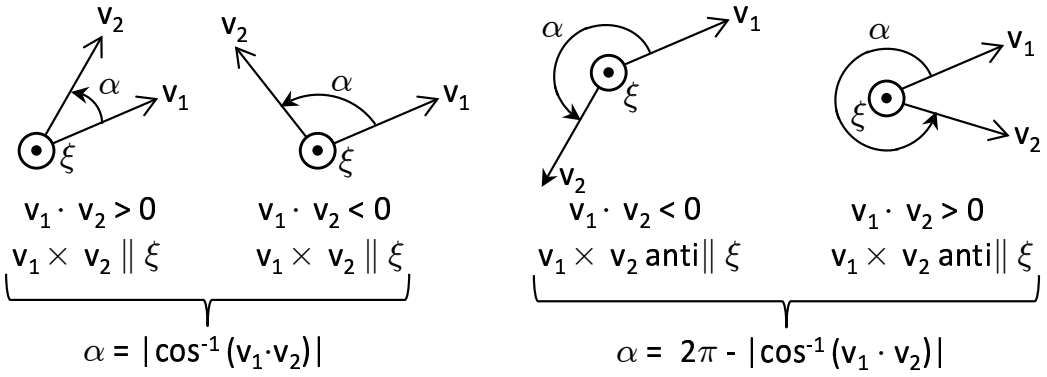

Figure 3.14: The four cases for determining Stokes vector rotation angles. $\mathbf{v}_1$ is the initial direction, which is rotated into the final direction $\mathbf{v}_2$ by a counterclockwise rotation about direction $\boldsymbol{\xi}$.

from the dot product of $\mathbf{h}_i$ and $\mathbf{p}$. However, there is some ambiguity in the angle obtained this way. This is resolved by recalling the choice of a positive rotation being counterclockwise when looking into the beam. In general, the rotation angle is the angle required to rotate the first vector $\mathbf{v}_1$ counterclockwise into the second vector $\mathbf{v}_2$ when looking into the rotation axis, which is vector $\mathbf{v}_3 = \mathbf{v}_1 \times \mathbf{v}_2$. In the present case, $\mathbf{v}_1 = \mathbf{h}_i$, $\mathbf{v}_2 = \mathbf{s}$, and $\mathbf{v}_3 = \boldsymbol{\xi}_i$ is the rotation axis. In general, four cases must be considered, as shown in Fig. 3.14. The rotation angle depends on whether $\mathbf{v}_1 \cdot \mathbf{v}_2$ is positive or negative, and on whether $\mathbf{v}_1 \times \mathbf{v}_2$ is parallel or antiparallel to the rotation axis. In the present case, $\mathbf{v}_1 \times \mathbf{v}_2 = \mathbf{h}_i \times \mathbf{s}$ is by construction parallel to $\boldsymbol{\xi}_i$, in which case

$$\alpha_i = |\cos^{-1}(\mathbf{v}_1 \cdot \mathbf{v}_2)| = |\cos^{-1}(\mathbf{h}_i \cdot \mathbf{s})|\,. \tag{3.53}$$

If $\mathbf{v}_1 \times \mathbf{v}_2$ is antiparallel to $\boldsymbol{\xi}$, then

$$\alpha_i = 2\pi - |\cos^{-1}(\mathbf{v}_1 \cdot \mathbf{v}_2)| = 2\pi - |\cos^{-1}(\mathbf{h}_i \cdot \mathbf{s})|\,. \tag{3.54}$$

Vector $\mathbf{v}_1 \times \mathbf{v}_2$ is parallel (antiparallel) to $\boldsymbol{\xi}$ if the z component of $\mathbf{v}_1 \times \mathbf{v}_2$ has the same (different) sign as the z component of $\boldsymbol{\xi}$.

When a ray finally leaves the hexagonal domain, its Stokes vector $\underline{S}_f$ must be specified in the final meridian plane. The final horizontal and vertical directions are obtained from choosing the first direction to be the final ray direction $\boldsymbol{\xi}_f (= \boldsymbol{\xi}_r$ or $\boldsymbol{\xi}_t)$ cross $\mathbf{z}$. Then, as above, $\mathbf{v}_f = \boldsymbol{\xi}_f \times \mathbf{h}_f$. The final rotation angle α_f, which carries the Stokes vector from the most recent wave facet $\mathbf{s}$ and $\mathbf{p}$ axes to the final $\mathbf{h}_f$ and $\mathbf{v}_f$ axes, is obtained by Eq. (3.53) or (3.54). The green sets of $\mathbf{h}_r, \mathbf{v}_r$ and $\mathbf{h}_t, \mathbf{v}_t$ vectors in the figure show the horizontal and vertical axes for the final meridian planes of the reflected and transmitted vectors. A green circumpolar dotted line makes clear that $\mathbf{h}_t$ is in the $-\boldsymbol{\varphi}$ direction, as is $\mathbf{h}_r$. The numerical values shown in the figure inset give the directions and rotation angles (all in degrees) for the vectors of this figure.

It should be remembered that all of the vectors seen in Fig. 3.13 apply at the point of intersection between a ray and a wave facet. In other words, all of the arrows in Fig. 3.13 should be drawn with their tails at the point of intersection, i.e. at the origin of the **x-y-z** coordinate system. This makes the figure hopelessly messy, but does make the rotations easier to visualize.

These rules for computing directions and rotation angles give all that is needed to carry Stokes vectors from one wave facet to another. Figure 3.15 shows an example of multiple scattering between surface wave facets. The left part of the figure shows a water-incident ray intersecting the first wave facet from below (thus the blue triangular wave facet hides the head of the $\boldsymbol{\xi}_i$ arrow as drawn). This scattering yields a reflected ray, which leaves the hexagonal domain. The Stokes vector for this ray is rotated from the scattering plane $\mathbf{s}_1, \mathbf{p}_1$ coordinates of the first facet to the final meridian plane horizontal and vertical $\mathbf{h}_r, \mathbf{v}_r$ system shown in the the left part of the figure. That rotation angle is given by α_r in the inset of the left figure.

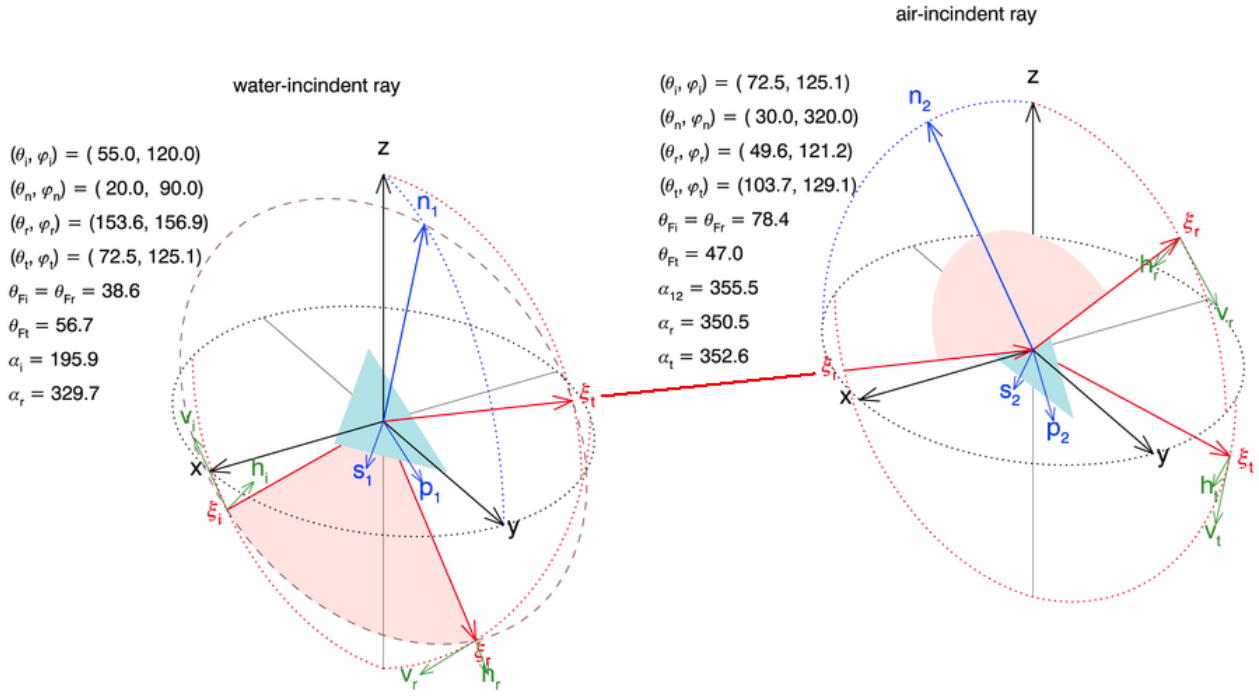

Figure 3.15: Illustration of multiple scattering by surface wave facets. A water-incident ray eventually yields one ray in the air and two in the water that leave the hexagonal domain and whose Stokes vectors are tallied. There is one intermediate ray from the first facet to the second that does not contribute directly to the final result.

The water-incident ray also yields a transmitted ray, which is heading slightly upward (at $\theta_t = 72.5$ deg, or 17.5 deg above the horizontal). Before leaving the hexagonal domain, that ray intersects another wave facet, now from the air side, as shown in the right part of the figure. The transmitted ray from the first facet, $\boldsymbol{\xi}_t$ of the left figure, becomes the incident ray for the second facet, $\boldsymbol{\xi}_i$ of the right figure. The Stokes vector $\underline{S}_t$ of the first transmitted ray is specified in the $\mathbf{s}_1, \mathbf{p}_1$ coordinates of the first facet. This Stokes vector must be rotated into the $\mathbf{s}_2, \mathbf{p}_2$ coordinates of the second facet. As in Eq. (3.52a), $\mathbf{s}_2 = \boldsymbol{\xi}_i \times \mathbf{n}_2 / |\boldsymbol{\xi}_i \times \mathbf{n}_2|$. For these particular vectors, $\mathbf{s}_1 \times \mathbf{s}_2$ is antiparallel to $\boldsymbol{\xi}_i$, so the rotation angle α_{12} that takes the Stokes vector from the first facet to the second is given by Eq. (3.54), where now this formula gives $\alpha_{12} = 2\pi - |\cos^{-1}(\mathbf{s}_1 \cdot \mathbf{s}_2)|$. The intersection of first transmitted ray with the second facet yields reflected and refracted rays.

Both of these rays leave the hexagonal domain, and their Stokes vectors are then rotated into their respective final meridian planes, as shown by the green $\mathbf{h}_r, \mathbf{v}_r$ and $\mathbf{h}_t, \mathbf{v}_t$ arrows of the right figure.

The next step of the ray tracing discussion is to specify exactly what is to be computed in the Monte Carlo simulation. Recall from the surface boundary conditions (2.21) and (2.22) that what is needed are four quad-averaged sea-surface transfer functions $\underline{r}_{aw}(r, s \to u, v)$, $\underline{r}_{wa}(r, s \to u, v)$, $\underline{t}_{aw}(r, s \to u, v)$, and $\underline{t}_{wa}(r, s \to u, v)$. These functions transform incident Stokes vectors $\underline{S}(r, s)$ into final Stokes vectors $\underline{S}(u, v)$. These transfer functions are inherent optical properties of the sea surface that describe its effect on incident light; they are independent of the incident light itself. The transfer functions are essentially scattering phase matrices that describe how the surface scatters light from any incident direction into any other direction, and as such they include the effects of the rotations from the incident meridian plane of $\underline{S}(r, s)$ to the final meridian plane of $\underline{S}(u, v)$. For the case of air-incident light being reflected back upward by the surface, and for multiple scattering between wave facets, we can write

$$\begin{aligned} \underline{S}(u, v) &= \underline{r}_{aw}(r, s \to u, v)\underline{S}(r, s) \\ &= \underline{R}(\alpha_f)\underline{M}(\psi_m)\underline{R}(\alpha_{m-1,m}) \cdots \underline{R}(\alpha_{1,2})\underline{M}(\psi_1)\underline{R}(\alpha_i)\underline{S}(r, s) \end{aligned} \tag{3.55}$$

Here

- α_i is the rotation angle that rotates the initial $\underline{S}(r, s)$ into the scattering plane of the first wave facet.
- $\underline{M}(\psi_1)$ is any of the four Fresnel reflectance and transmittance functions $\underline{R}_{aw}, \underline{R}_{wa}, \underline{T}_{aw}$, or $\underline{T}_{wa}$ for a level water surface, as given in §3.1.
- ψ_1 represents the scattering angle for the incident and reflected or transmitted directions used in the Fresnel function. ψ_1 thus represents a θ_i, θ_r or θ_i, θ_t pair depending on whether the incident ray is reflected or transmitted at facet 1.
- $\underline{R}(\alpha_{1,2})$ is the rotation matrix that takes the Stokes vector $\underline{M}(\psi_1)\underline{R}(\alpha_i)\underline{S}(r, s)$ from the scattering plane of the first wave facet to the scattering plane of the second wave facet.
- $\underline{R}(\alpha_{m-1,m})$ carries the current Stokes vector from wave facet $m-1$ to the last facet m.
- $\underline{M}(\psi_m)$ is the Fresnel reflection or transmission function for the last wave facet.
- $\underline{R}(\alpha_f)$ is the rotation matrix that carries the Stokes vector from the scattering plane of the last wave facet to the final meridian plane associated with $\underline{S}(u, v)$.

For given quads $Q(r, s)$ and $Q(u, v)$, the surface transfer function is a product of rotation matrices and Fresnel reflectance and transmittance matrices. These rotation and Fresnel matrices are determined by the particular wave facets intercepted by the ray path connecting $Q(r, s)$ and $Q(u, v)$. Note that a single air-incident ray usually contributes to both $\underline{r}_{aw}$ and $\underline{t}_{aw}$, since the reflected and refracted daughter rays can leave the hexagonal domain in any direction. Likewise, a water-incident ray usually contributes to both $\underline{r}_{wa}$ and $\underline{t}_{wa}$. Thus in the ray tracing described next, we need to keep track of four accumulating weight arrays, $\underline{\mathcal{W}}^{(\mathfrak{s})}_{raw}(r, s \to u, v)$ and three others. The subscript on $\underline{\mathcal{W}}$ specifies which of the four weight arrays is referenced: $\underline{\mathcal{W}}_{raw}$ for the weight array

leading eventually to $\underline{r}_{aw}$, and $\underline{\mathcal{W}}_{taw}$, $\underline{\mathcal{W}}_{rwa}$, or $\underline{\mathcal{W}}_{twa}$ for the other four functions. The superscript index $(\mathfrak{s})$ labels that the weight array is for the $\mathfrak{s}^{th}$ surface realization. As seen next, a particular ray has associated with it a single weight array $\underline{\mathcal{W}}$, which will be tallied to one of the four accumulating weight arrays $\underline{\mathcal{W}}^{(\mathfrak{s})}_{raw}(r, s \to u, v)$, etc.

The flow chart of Fig. 3.12 outlines the general procedure for Monte Carlo computation of $\underline{r}_{aw}(r, s \to u, v)$ and the other three surface functions. The details of the calculations for the $\mathfrak{s}^{th}$ surface realization and a single initial ray from one quad $Q(r, s)$ are as follows:

1. Choose a random point within $Q(r, s)$ and aim a ray $\boldsymbol{\xi}_i$ from that point to a particular target point $\mathbf{t}$ near the center of the hexagonal grid used to define the current sea surface realization.

2. Compute the point $\mathbf{p}_i$ where the initial ray enters the hexagonal domain as described in §B.1.

3. Compute the initial horizontal direction $\mathbf{h}_i$ from Eq. (3.51a), as is used to specify Stokes vectors in the initial meridian plane.

4. Initialize the weight matrix $\underline{\mathcal{W}}_i$ to the 4×4 identity matrix: $\underline{\mathcal{W}}_i = \underline{I}_{4\times4}$.

5. Push the initial ray information $(\mathbf{p}_i, \boldsymbol{\xi}_i, \mathbf{h}_i, \underline{\mathcal{W}}_i)$ into the stack. This completes the ray initialization.

6. Pull the next available ray $(\mathbf{p}', \boldsymbol{\xi}', \mathbf{s}', \underline{\mathcal{W}}')$ from the stack and trace it to completion. (If the ray pulled from the stack is an initial ray, then $\mathbf{s}'$ as pulled will the value of $\mathbf{h}_i$. If the ray pulled has already intersected a facet, then the pulled $\mathbf{s}'$ will be the $\mathbf{s}$ of the previous facet.)

   (a) If the ray intersects a wave facet:

      i. Compute the direction $\mathbf{s}$ perpendicular to the scattering plane for the intersection of the ray with the facet, using Eq. (3.52a).

      ii. Compute the rotation angle α using Eq. (3.53) or (3.54) with $\mathbf{v}_1 = \mathbf{s}'$ and $\mathbf{v}_2 = \mathbf{s}$.

      iii. Update the weight using $\underline{\mathcal{W}} = \underline{R}(\alpha)\underline{\mathcal{W}}'$. The $\underline{R}(\alpha)$ rotation carries the Stokes vector from its previous reference plane into the scattering plane of the wave facet.

      iv. Determine the directions of the reflected ($\boldsymbol{\xi}_r$) and transmitted ($\boldsymbol{\xi}_t$, if there is one) rays at the point $\mathbf{p}$ of intersection.

      v. For the reflected ray, multiply the weight $\underline{\mathcal{W}}$ by the Fresnel reflectance function $\underline{R}_{\mathrm{F}} = \underline{R}_{aw}$ or $\underline{R}_{wa}$, depending on whether the initial ray is air-incident or water-incident. That is, set the weight for the reflected ray to be $\underline{\mathcal{W}}_r = \underline{R}_{\mathrm{F}}\underline{\mathcal{W}}$. The angles of incidence and transmission for use in $\underline{R}_{\mathrm{F}}$ are determined from the dot product of the incident vector and the normal to the wave facet. Push the reflected ray information $(\mathbf{p}, \boldsymbol{\xi}_r, \mathbf{s}, \underline{\mathcal{W}}_r)$ into the stack to await further processing.

      vi. For the transmitted ray (if there is one), multiply the weight $\underline{\mathcal{W}}$ by the Fresnel transmittance function $\underline{T}_{\mathrm{F}} = \underline{T}_{aw}$ or $\underline{T}_{wa}$, depending on whether the initial ray is air-incident or water-incident. That is, set the weight for the transmitted ray to be $\underline{\mathcal{W}}_t = \underline{T}_{\mathrm{F}}\underline{\mathcal{W}}$. Push the transmitted ray information $(\mathbf{p}, \boldsymbol{\xi}_t, \mathbf{s}, \underline{\mathcal{W}}_t)$ into the stack to await further processing.

   (b) If the ray leaves the hexagonal domain:

i. Determine the quad $Q(u,v)$ receiving the final ray. The initial and final quads for this ray path determine which of the four weight arrays $\underline{\mathcal{W}}^{(\mathfrak{s})}_{raw}(r,s\rightarrow u,v)$ etc. will receive the ray's final weight array.

ii. Compute the final horizontal direction $\mathbf{h}_f = \boldsymbol{\xi}' \times \mathbf{z}$.

iii. Compute the final rotation angle α_f that takes the Stokes vector into the final meridian plane.

iv. Update the ray's weight array to its final value: $\underline{\mathcal{W}}_f = \underline{R}(\alpha_f)\underline{\mathcal{W}}'$.

v. Add the final weight for this ray to the accumulating weight array that connects quads $Q(r,s)$ and $Q(u,v)$, e.g. $\underline{\mathcal{W}}^{(\mathfrak{s})}_{raw}(r,s\rightarrow u,v) := \underline{\mathcal{W}}^{(\mathfrak{s})}_{raw}(r,s\rightarrow u,v) + \underline{\mathcal{W}}_f$.

7. If the stack is not empty, return to Step 6 and begin processing the next ray. If the stack is empty, return to Step 1 and initialize a new ray from a different quad. Repeat this cycle until a ray has been traced from a random direction within each quad, for the given surface realization.

The above steps trace a ray from a random direction within each quad, for a given surface realization. Note that a single initial ray in Step 1 generally contributes to the accumulating weights for two or more different final quads. The exception occurs if the initial ray is totally internally reflected and that reflected ray leaves the hexagonal domain without further intersections with the surface. In that case, only one final quad receives the entire ray weight.

This entire process is repeated for a large number $\mathfrak{S}$ of sea surface realizations; typically $\mathfrak{S} = 10^5$ or more. After tracing rays from each quad for each surface realization, we compute the ensemble averages of the four weight arrays as

$$\langle \underline{\mathcal{W}}_{raw}(r,s\rightarrow u,v)\rangle = \frac{1}{\mathfrak{S}}\sum_{\mathfrak{s}=1}^{\mathfrak{S}} \underline{\mathcal{W}}^{(\mathfrak{s})}_{raw}(r,s\rightarrow u,v)\,, \tag{3.56}$$

with corresponding equations for the other three weight arrays. (Note: Eq. (3.56) represents the conceptual process. However, the code does not store separate arrays for each surface realization; the four arrays are just added to until the prescribed number of surface simulations has been reached. Then the final value of $\underline{\mathcal{W}}^{(\mathfrak{s})}_{raw}(r,s\rightarrow u,v)$ is divided by $\mathfrak{S}$.)

The ray tracing calculations just described give four ensemble averages $\langle \underline{\mathcal{W}}_{raw}(r,s\rightarrow u,v)\rangle$, etc. These quantities are the averages of the product $\underline{R}(\alpha_f)\underline{M}(\psi_m)\cdots\underline{M}(\psi_1)\underline{R}(\alpha_i)$, shown in Eq. (3.55), for many different ray paths connecting quads $Q(r,s)$ and $Q(u,v)$. One final step remains to convert these ensemble averages into the surface radiance transfer functions seen in the boundary conditions.

Recall that the first element of a Stokes vector $\underline{S} = [I,Q,U,V]^T$ gives the total radiance without regard to its state of polarization. The I term thus represents the total energy of the light ray (within the solid angle of the quad, measured on a surface normal to the direction of propagation). The (1,1) elements of the Fresnel reflectance and transmittance matrices show how much of this energy is reflected and transmitted as a function of the incident angle of the ray with a facet normal in each ray-facet interaction. Conservation of energy of course requires that the reflected plus transmitted energy equal the incident energy in each interaction. The (1,1) elements of the rotation matrices (1.14) are always 1 and thus do not affect the I element of Stokes vectors. The

weight array for each initial ray is the identity matrix, so the (1,1) element is 1. At each ray-facet interaction, this initial value of 1 will be reduced by the value of the (1,1) element of the Fresnel reflectance or transmittance. Thus we can see that *the final value of the (1,1) element of the ensemble-averaged weight array* $\langle \underline{\mathcal{W}}_{raw}(r, s \to u, v) \rangle$ *gives the fraction of the initial total energy of rays from quad* $Q(r, s)$ *that eventually reaches quad* $Q(u, v)$. A similar interpretation applies to the other matrix elements. Thus, for example, the (2,3) element of $\underline{\mathcal{W}}$ gives the fraction of energy coming from quad $Q(r, s)$ in a horizontal ($Q > 0$) or vertical ($Q < 0$) polarization state in the incident meridian plane then ends up in quad $Q(u, v)$ in a polarization state that is at +45 deg ($U > 0$) or -45 deg ($U < 0$) in the final meridian plane. The non-dimensional ensemble averages can therefore be interpreted as energy (or power) transfer functions.

Conservation of total energy at the sea surface therefore can be written in terms of the (1,1) elements of the $\underline{\mathcal{W}}$ arrays as

$$\sum_{\substack{u \\ u,v \in \Xi_u}} \sum_{v} \langle \underline{\mathcal{W}}_{raw}(r, s \to u, v) \rangle_{1,1} + \sum_{\substack{u \\ u,v \in \Xi_d}} \sum_{v} \langle \underline{\mathcal{W}}_{taw}(r, s \to u, v) \rangle_{1,1} = 1 \,, \tag{3.57}$$

for every $Q(r, s)$ in Ξ_d. The notation here means that downwelling energy coming from any given incident quad is either reflected back upward into some quad $Q(u, v)$ in Ξ_u (the first term on the left) or transmitted downward into some quad $Q(u, v)$ in Ξ_d (the second term) without loss. The corresponding conservation-of-energy equation for upwelling incident energy is

$$\sum_{\substack{u \\ u,v \in \Xi_d}} \sum_{v} \langle \underline{\mathcal{W}}_{rwa}(r, s \to u, v) \rangle_{1,1} + \sum_{\substack{u \\ u,v \in \Xi_u}} \sum_{v} \langle \underline{\mathcal{W}}_{twa}(r, s \to u, v) \rangle_{1,1} = 1 \,, \tag{3.58}$$

for every $Q(r, s) = Q_{rs}$ in Ξ_u. These equations can be compared with the surface boundary conditions (2.21) and (2.22),

$$\begin{aligned} \underline{S}(a, u, v) = & \sum_{\substack{r \\ r,s \in \Xi_d}} \sum_{s} \underline{r}_{aw}(r, s \to u, v)\, \underline{S}(a, r, s) \\ & + \sum_{\substack{r \\ r,s \in \Xi_u}} \sum_{s} \underline{t}_{wa}(r, s \to u, v)\, \underline{S}(w, r, s) \quad \text{for (u, v)} \in \Xi_{\text{u}} \end{aligned} \tag{3.59}$$

and

$$\begin{aligned} \underline{S}(w, u, v) = & \sum_{\substack{r \\ r,s \in \Xi_u}} \sum_{s} \underline{r}_{wa}(r, s \to u, v)\, \underline{S}(w, r, s) \\ & + \sum_{\substack{r \\ r,s \in \Xi_d}} \sum_{s} \underline{t}_{aw}(r, s \to u, v)\, \underline{S}(a, r, s) \quad \text{for (u, v)} \in \Xi_{\text{d}} \,. \end{aligned} \tag{3.60}$$

The surface boundary conditions show where the radiance (energy) comes from for a given output direction $Q(u, v)$. The conservation-of-energy equations show where the energy goes for a given input direction $Q(r, s)$.

Equations (3.57) to (3.60) hold for any input quad $Q(r, s)$. The incident radiant power per unit horizontal area of the mean sea surface is the plane irradiance E. The downwelling plane irradiance onto the air side of the sea surface (depth index a) due to a single quad $Q(r, s) \in \Xi_d$ is

$$E_d(a, r, s) = I(a, r, s)\, |\mu_r|\, \Omega_{rs} \,. \tag{3.61}$$

The upwelling irradiance in quad $Q(u,v) \in \Xi_u$ generated when the sea surface reflects this downwelling irradiance is

$$E_u(a,u,v) = E_d(a,r,s)\,\langle \underline{\mathcal{W}}_{raw}(r,s \to u,v)\rangle_{1,1}\,, \tag{3.62}$$

since, as we have seen, $\langle \underline{\mathcal{W}}_{raw}(r,s \to u,v)\rangle_{1,1}$ is the fraction of the total energy in $Q(r,s)$ that is reflected into $Q(u,v)$. Writing $E_u(a,u,v)$ in terms of the radiance, as in Eq. (3.61), the last equation becomes

$$I(a,u,v)|\mu_u|\,\Omega_{uv} = I(a,r,s)|\mu_r|\,\Omega_{rs}\langle \underline{\mathcal{W}}_{raw}(r,s \to u,v)\rangle_{1,1}\,, \tag{3.63}$$

or

$$I(a,u,v) = I(a,r,s)\,\left[\frac{|\mu_r|\,\Omega_{rs}\langle \underline{\mathcal{W}}_{raw}(r,s \to u,v)\rangle_{1,1}}{|\mu_u|\,\Omega_{uv}}\right]. \tag{3.64}$$

Now recall the surface boundary condition (3.59). This equation holds even if only one particular input quad $Q(r,s)$ has a nonzero Stokes vector and all other quads are dark, as is the case for Eq. (3.64). Equation (3.59) for the first element of the Stokes vector then reduces to

$$I(a,u,v) = I(a,r,s)\,\underline{r}_{aw}(r,s \to u,v)_{1,1}\,. \tag{3.65}$$

Comparing Eqs. (3.64) and (3.65) immediately yields the connection between the (1,1) element of the surface radiance transfer function $\underline{r}_{wa}$ and the (1,1) element of the ensemble-averaged energy transfer function $\underline{\mathcal{W}}_{raw}$:

$$\underline{r}_{aw}(r,s \to u,v)_{1,1} = \langle \underline{\mathcal{W}}_{raw}(r,s \to u,v)\rangle_{1,1}\,\frac{|\mu_r|\,\Omega_{rs}}{|\mu_u|\,\Omega_{uv}}\,. \tag{3.66}$$

Thus the *energy* transfer function $\underline{\mathcal{W}}_{raw}$ must be scaled by the cosines of the incident and reflected directions and the solid angles of the incident and reflected quads in order to obtain the corresponding *radiance* transfer function $\underline{r}_{aw}$.

Applying the same scaling to all elements of the ensemble-averaged energy transfer functions gives the needed final part of the computation of the radiance transfer functions (L&W 4.74):

$$\underline{r}_{aw}(r,s \to u,v) = \langle \underline{\mathcal{W}}_{raw}(r,s \to u,v)\rangle\,\frac{|\mu_r|\,\Omega_{rs}}{|\mu_u|\,\Omega_{uv}}\,, \tag{3.67a}$$

$$\underline{t}_{aw}(r,s \to u,v) = \langle \underline{\mathcal{W}}_{taw}(r,s \to u,v)\rangle\,\frac{|\mu_r|\,\Omega_{rs}}{|\mu_u|\,\Omega_{uv}}\,, \tag{3.67b}$$

$$\underline{r}_{wa}(r,s \to u,v) = \langle \underline{\mathcal{W}}_{rwa}(r,s \to u,v)\rangle\,\frac{|\mu_r|\,\Omega_{rs}}{|\mu_u|\,\Omega_{uv}}\,, \tag{3.67c}$$

$$\underline{t}_{wa}(r,s \to u,v) = \langle \underline{\mathcal{W}}_{twa}(r,s \to u,v)\rangle\,\frac{|\mu_r|\,\Omega_{rs}}{|\mu_u|\,\Omega_{uv}}\,. \tag{3.67d}$$

Evaluation of these equations completes the Monte Carlo simulation of the sea-surface radiance transfer functions needed for the surface boundary conditions.

Equations 3.67 can be used in Eqs. (3.57) and (3.58) to express conservation of total energy at the sea surface in terms of the radiance transfer functions:

$$\frac{1}{|\mu_r|\,\Omega_{rs}}\sum_{\substack{u\\u,v\in\Xi_u}}\sum_{v}\underline{r}_{aw}(r,s \to u,v)_{1,1}\,|\mu_u|\,\Omega_{uv} + \frac{1}{|\mu_r|\,\Omega_{rs}}\sum_{\substack{u\\u,v\in\Xi_d}}\sum_{v}\underline{t}_{aw}(r,s \to u,v)_{1,1}\,|\mu_u|\,\Omega_{uv} = 1\,, \tag{3.68}$$

for every $Q(r,s)$ in Ξ_d. A corresponding equation holds for upward incident radiance, as in Eq. (3.58).

Two important observations can be made about these calculations. First, the n^2 law for radiance is built into the radiance transfer functions photon by photon during the Monte Carlo ray tracing, just as happens in nature. Second, the ray-by-ray conservation of energy in the Monte Carlo simulations guarantees global (hemisphere-wide) conservation of energy in the radiance reflectance and transmittance functions.

Hovenier (1969) derived 20 symmetry relations (his relations a through t) for finitely thick layers of multiply scattering particles whose phase matrices have the symmetry of Eq. (1.36). These relations are similar to those seen for phase matrices in Eqs. (1.35a) to (1.37d). The Fresnel reflection and scattering matrices of Eqs. (3.2) and (3.3) obey the symmetry of Eq. (1.36). It therefore seems that the surface reflectance and transmission functions for a wind-blown surface should obey the Hovenier symmetry relations as well, with multiple scattering by surface wave facets corresponding to multiple scattering by particles in a volume. Further investigation of these symmetries by numerical simulation of wind-blown surfaces is warranted.

CHAPTER 4

Solution of the VRTE within the Water

There are many numerical techniques for solving the radiative transfer equation. The one used in both HydroLight and HydroPol is invariant imbedding. The details of invariant imbedding theory as applied to the scalar RTE are given in Chapter 8 of *Light and Water* (Mobley, 1994). The scalar technique used in HydroLight is extended here to the vector RTE for use in HydroPol.

The RTE and its boundary conditions constitute a linear (in the radiances) two-point boundary value problem. That is, the linear RTE must be solved subject to boundary conditions at both the sea surface and bottom. *The essence of invariant imbedding is that it converts a linear two-point boundary value problem into a pair of non-linear initial value problems.* In the present case, one of these initial value problems is solved starting with a boundary value at the sea surface and integrating downward, and the other is solved starting with a boundary value at the sea bottom and integrating upward.

Invariant imbedding is mathematically complicated and consequently tedious to program, but the payoff is that it is both much faster in run time than the mathematically simpler (and therefore widely used) Monte Carlo technique and more general in the allowed inputs (such as allowing arbitrary depth dependence of the IOPs) than some other techniques. In particular, the run time for invariant imbedding is linearly proportional to the optical depth, whereas run times increase exponentially with optical depth for Monte Carlo solutions. Invariant imbedding allows for arbitrary depth dependence of the IOPs, whereas other techniques (e.g., discrete ordinates and adding-doubling methods) build up the water column as a stack of homogeneous layers, and run time is proportional to the number of layers used.

4.1 Physical Space vs. Fourier Space

Invariant imbedding theory can be used to solve directly for the discretized Stokes vectors $\underline{S}(k,u,v,j)$. There is an argument for doing this, namely that the mathematical development below is much easier and the quantities involved all have simple interpretations, e.g., as physical reflectances and transmittances that are bounded by 0 and 1. However, as will be seen in §4.5.4, a physical-space solution algorithm requires the simultaneous solution of a set of nonlinear ordinary differential equations (ODEs) of approximate size $128M^2N$ (ignoring internal source terms and special cases

for the polar caps). For the current quad partitioning with $M = 10, N = 12$, this gives a system of 153,600 equations. That is a formidable numerical problem.

As shown in Appendix A, a discrete function $f(v)$ of azimuthal angle $\phi_v, v = 1, ..., 2N$, can be represented as sines and cosines via

$$f(v) = \sum_{\ell=0}^{N} \left[\hat{f}_1(\ell) \cos(\ell\phi_v) + \hat{f}_2(\ell) \sin(\ell\phi_v) \right] .$$

This results in $N + 1$ generally non-zero cosine amplitudes $\hat{f}_1(\ell)$ and $N - 1$ sine amplitudes $\hat{f}_2(\ell)$. When applied to the VRTE, Fourier analysis leads to the solution of $N + 1$ sets of ODEs of size $64M^2$ equations for cosine amplitudes, and $N - 1$ sets of $64M^2$ equations for sine amplitudes, for the same total of $128M^2N$ equations solved as for the physical-space formulation. However, with Fourier analysis the ODE solutions are made as a sequence of $N+1$ independent "small" problems, rather than as one "large" problem.

HydroPol, like HydroLight, therefore uses a Fourier analysis of the VRTE in azimuthal angle to minimize computer run times. The penalty paid for this increase in numerical efficiency is a corresponding increase in the complexity of the mathematical development and computer programming.

4.2 Recasting the VRTE as Upward and Downward Equations

Recall the discretized VRTE of Eq. (2.20):

$$\begin{aligned} \mu_u \frac{d\underline{S}(\zeta, u, v, j)}{d\zeta} &= -\underline{S}(\zeta, u, v, j) \\ &\quad + \omega_o(\zeta, j) \sum_r \sum_s \underline{\tilde{P}}(\zeta; r, s \to u, v; j)\, \underline{S}(\zeta, r, s, j) \\ &\quad + \underline{\Sigma}(\zeta, u, v, j) \ , \end{aligned} \tag{4.1}$$

where $u = 1, ..., 2M$ and $v = 1, ..., 2N$ to cover all directions (with special cases for the polar caps).

Invariant imbedding requires that the VRTE be written as a pair of equations, one for upward directions and one for downward directions. In addition, the sea surface boundary conditions are formulated in terms of transmission and reflection operators that transfer radiance back and forth across the sea surface. As was shown in Fig. 1.2, the sea surface is described mathematically as a nonabsorbing "layer" of zero physical thickness that represents a discontinuity in the index of refraction and therefore causes scattering (reflection and refraction) of light incident onto the surface from above or below. A depth value of $\zeta = a$ denotes a location in the air just above the mean sea surface, and a depth value of $\zeta = w$ denotes a location at depth 0, but in the water just below the mean sea surface. An opaque bottom at depth $\zeta = m$ is defined by its reflectance properties. Figure 4.1 shows a redrawn version of Fig. 1.2. As in Figs. 1.2 and 2.1, depth is measured positive downward from 0 at the mean sea surface. Because direction $+\zeta$ is downward, a superscript + will denote downwelling radiances (photons heading downward); superscript − will denote upwelling radiances (photons heading upward). Thus in the figure, the downward arrow labeled $\underline{S}^+(a)$ represents downwelling radiance incident onto the sea surface from above, and $\underline{S}^-(a)$ represents upwelling radiance just above the surface, i.e., water-leaving radiance. $\underline{S}^+(w)$ represents radiance that has been transmitted through the surface and is incident onto the water below. The

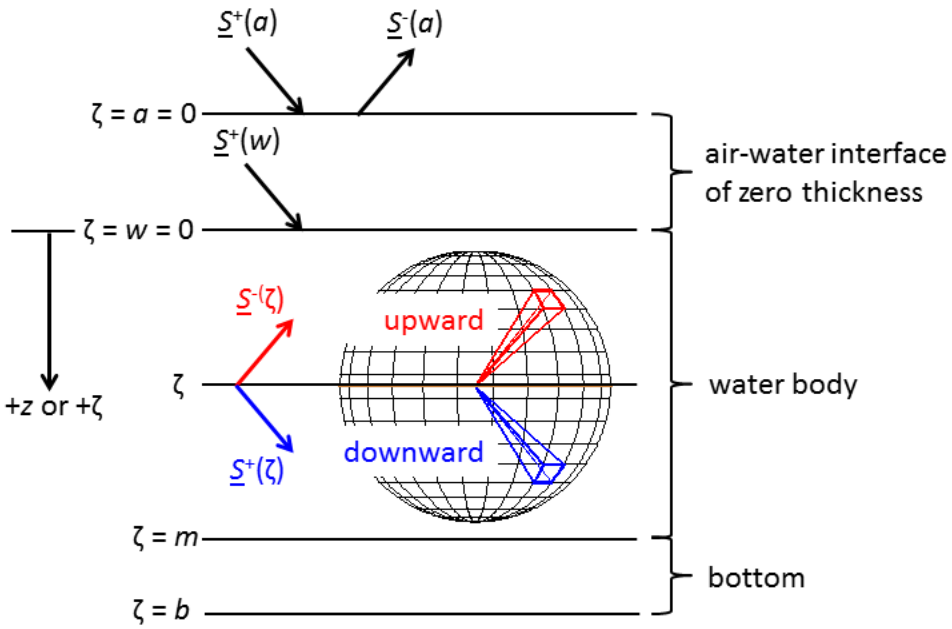

Figure 4.1: The depth coordinate system used for invariant imbedding. Upwelling and downwelling directions are indicated, along with a few representative Stokes vectors.

sphere of directions at depth ζ highlights a quad containing an upwelling radiance $\underline{S}^-(\zeta)$ (red quad and red arrow). The blue quad and blue arrow indicate downwelling direction and radiance.

Recall also that $\mu_u = \cos(\theta_u)$ in Eq. (4.1) ranges from -1 when $\theta = 180$ (photons heading straight up) to $+1$ when $\theta = 0$ (photons heading straight down). It will be convenient to redefine μ_u as a positive number and incorporate the negative μ values for upwelling radiance via an explicit minus sign in the VRTE. With these conventions, and dropping the wavelength argument j for simplicity of notation, the VRTE can be written as a pair of equations:

$$\begin{aligned} \mp\mu_u \frac{d\underline{S}^{\mp}(\zeta,u,v)}{d\zeta} &= -\underline{S}^{\mp}(\zeta,u,v) \\ &\quad + \omega_o(\zeta) \sum_r \sum_s \underline{\tilde{P}}^{\pm}(\zeta; r,s \to u,v)\, \underline{S}^-(\zeta,r,s) \\ &\quad + \omega_o(\zeta) \sum_r \sum_s \underline{\tilde{P}}^{\mp}(\zeta; r,s \to u,v)\, \underline{S}^+(\zeta,r,s) \\ &\quad + \underline{\Sigma}^{\mp}(\zeta,u,v)\ . \end{aligned} \tag{4.2}$$

Now $u = 1, ..., M$ in each hemisphere of directions (quad M is the polar cap), $v = 1, ..., 2N$, and $\mu_u > 0$. Note that now there are two radiances with the same (u, v) quad indices, but one is in the upper hemisphere ($-$ superscript) and one is in the lower ($+$ superscript). The $+$ superscript on the phase matrix $\underline{\tilde{P}}^+(\zeta; r, s \to u, v)$ denotes scattering between quads $Q(r, s)$ and $Q(u, v)$ that lie in the same hemisphere. The $-$ superscript on $\underline{\tilde{P}}^-(\zeta; r, s \to u, v)$ denotes scattering between quads in opposite hemispheres. This is illustrated in Fig. 4.2. Scattering from the green quad $Q(r, s)$ to the red quad $Q(u, v)$ is described by $\underline{\tilde{P}}^+(\zeta; r, s \to u, v)$; scattering from the green to the blue quad

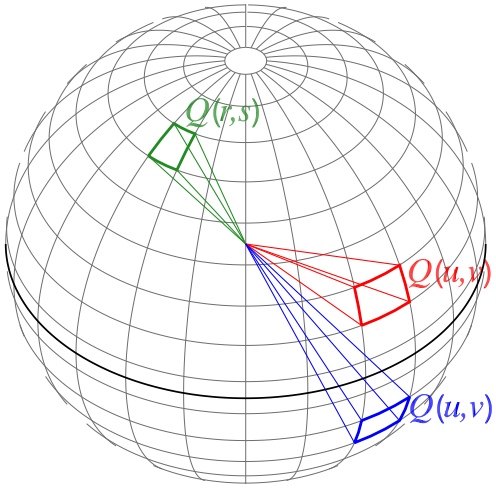

Figure 4.2: Illustration of scattering within and between hemispheres of directions. Scattering within a hemisphere, e.g. from the green quad $Q(r,s)$ to the red quad $Q(u,v)$, is described by $\tilde{P}^{+}(r,s,u,v)$. Scattering between hemispheres, e.g. from the green quad $Q(r,s)$ to the blue quad $Q(u,v)$, is described by $\tilde{P}^{-}(r,s,u,v)$.

has a different scattering angle ψ and is described by $\underline{\tilde{P}}^{-}(\zeta;r,s\rightarrow u,v)$. Taking the top sign in Eq. (4.2) gives the VRTE for upwelling radiances $\underline{S}^{-}(\zeta,u,v)$, and the bottom sign gives the equation for downwelling radiances $\underline{S}^{+}(\zeta,u,v)$.

4.3 Fourier Decomposition of the Upward and Downward VRT Equations

The next step is to Fourier decompose the upward and downward pair of VRTE of equations. As was noted in §1.4.2, some terms of the phase matrix depend on $\cos(\phi_v-\phi_s)$ and some depend on $\sin(\phi_v-\phi_s)$. Rather than doing different Fourier expansions for different phase matrix elements and keeping track of which is which, it is notationally convenient to do a full Fourier expansion for each matrix element. However, when performing actual computations, one half of the calculations can be avoided by setting either the cosine or the sine amplitudes to zero, depending on which matrix element is being considered. Appendix A contains for reference the various formulas for Fourier decomposition of discrete functions of one or two azimuthal angles, as required here.

The Stokes vectors are functions of a single azimuthal angle and are decomposed as (Eq. A.6)

$$\underline{S}^{\pm}(\zeta,u,v)=\sum_{\ell=0}^{N}\left[\underline{\hat{S}}_1^{\pm}(\zeta,u|\ell)\cos(\ell\phi_v)+\underline{\hat{S}}_2^{\pm}(\zeta,u|\ell)\sin(\ell\phi_v)\right], \tag{4.3}$$

where $v=1,2,...,2N$. The vertical bar in the $(\zeta,u|\ell)$ argument list is used to separate the physical variables from the Fourier mode variable ℓ. The cosine and sine amplitudes are computed via Eqs. (A.7) and (A.9), respectively. A similar equation holds for the source term $\underline{\Sigma}$. In all amplitude variables here and below, subscript 1 denotes a cosine amplitude, and subscript 2 a sine amplitude.

As noted in §1.4.2, the elements of the phase matrix depend on either $\cos(\phi_v-\phi_s)$ or $\sin(\phi_v-\phi_s)$. Rather than keep track of which element has which dependence, it is notationally convenient to

expand each element of $\underline{P}^{\pm}(\zeta, r, s \to u, v)$ as

$$\underline{P}^{\pm}(\zeta, r, s \to u, v) = \sum_{k=0}^{N} \left[\underline{\hat{P}}_1^{\pm}(\zeta, r, u|k) \cos[k(\phi_v - \phi_s)] + \underline{\hat{P}}_2^{\pm}(\zeta, r, u|k) \sin[k(\phi_v - \phi_s)] \right] \tag{4.4}$$

However, in the computer code, either $\hat{P}^{\pm}_{1_{i,j}}$ or $\hat{P}^{\pm}_{2_{i,j}}$ will be zero for a particular $P^{\pm}_{i,j}$ element.

The non-zero amplitudes in Eq. (4.4) are computed by Eqs. (A.12) and (A.14). However, polar caps must always be treated as special cases (L&W 8.34(i)-8.34(iv)). Recall also that $s = 1$ or $\phi' = 0$ was chosen as the reference point to anchor the azimuthal dependence on $\phi - \phi'$. Then depending on which element of $\underline{P}^{\pm}(\zeta, r, s \to u, v)$ is being Fourier analyzed, either the cosine or sine formulas are used as follows (with the other amplitude set to zero):

(i) **Quad-to-quad scattering**: $u, r = 1, 2, ..., M-1$. Equations Eqs. (A.12) and (A.14) hold as written. Thus for given r, u values,

$$\underline{\hat{P}}_1^{\pm}(r, u|k) = \frac{1}{\epsilon_k} \sum_{v=1}^{2N} \underline{P}^{\pm}(r, s = 1 \to u, v) \cos(k\phi_v) \qquad \text{for } k = 0, ..., N \tag{4.5a}$$

$$\underline{\hat{P}}_2^{\pm}(r, u|k) = \frac{1}{\gamma_k} \sum_{v=1}^{2N} \underline{P}^{\pm}(r, s = 1 \to u, v) \sin(k\phi_v) \qquad \text{for } k = 1, ..., N-1 \tag{4.5b}$$

$$\underline{\hat{P}}_2^{\pm}(r, u|k) = 0 \qquad \text{for } k = 0 \text{ or N} \tag{4.5c}$$

where ϵ_ℓ and γ_ℓ are defined by Eqs. (A.8) and (A.10).

(ii) **Polar cap-to-quad scattering**: $r = M; u = 1, 2, ..., M-1$.

$$\underline{\hat{P}}_1^{\pm}(M, u|k) = \underline{P}^{\pm}(M, \cdot \to u, v) \qquad \text{for } k = 0 \tag{4.5d}$$

$$\underline{\hat{P}}_1^{\pm}(M, u|k) = 0 \qquad \text{for } k = 1, ..., N \tag{4.5e}$$

$$\underline{\hat{P}}_2^{\pm}(M, u|k) = 0 \qquad \text{for } k = 0, ..., N \tag{4.5f}$$

(iii) **Quad-to-Polar cap scattering**: $r = 1, 2, ..., M-1; u = M$.

$$\underline{\hat{P}}_1^{\pm}(r, M|k) = \underline{P}^{\pm}(r, s = 1 \to M, \cdot) \qquad \text{for } k = 0 \tag{4.5g}$$

$$\underline{\hat{P}}_1^{\pm}(r, M|k) = 0 \qquad \text{for } k = 1, ..., N \tag{4.5h}$$

$$\underline{\hat{P}}_2^{\pm}(r, M|k) = 0 \qquad \text{for } k = 0, ..., N \tag{4.5i}$$

(iv) **Polar cap-to-Polar cap scattering**: $r = u = M$.

$$\underline{\hat{P}}_1^{\pm}(M, M|k) = \underline{P}^{\pm}(M, \cdot \to M, \cdot) \qquad \text{for } k = 0 \tag{4.5j}$$

$$\underline{\hat{P}}_1^{\pm}(M, M|k) = 0 \qquad \text{for } k = 1, ..., N \tag{4.5k}$$

$$\underline{\hat{P}}_2^{\pm}(M, M|k) = 0 \qquad \text{for } k = 0, ..., N \tag{4.5l}$$

The dot argument for s or v for polar caps is a reminder that polar caps have no ϕ' or ϕ dependence. In the computer code, the polar cap values can be stored in the $r = M, s = 1$ or $u = M, v = 1$ array locations, with the remaining array elements for $r, u = M$ being unused and set to zero.

Inserting these Fourier decompositions into Eq. 4.2 gives (L&W 8.35)

$$
\begin{aligned}
&\mp\mu_u \frac{d}{d\zeta} \sum_{\ell=0}^{N} \left[\underline{\hat{S}}_1^{\mp}(\zeta, u|\ell) \cos(\ell\phi_v) + \underline{\hat{S}}_2^{\mp}(\zeta, u|\ell) \sin(\ell\phi_v) \right] \\
&= -\sum_{\ell=0}^{N} \left[\underline{\hat{S}}_1^{\mp}(\zeta, u|\ell) \cos(\ell\phi_v) + \underline{\hat{S}}_2^{\mp}(\zeta, u|\ell) \sin(\ell\phi_v) \right] \\
&+ \omega_o(\zeta) \sum_{r=1}^{M-1} \sum_{s=1}^{2N} \left\{ \sum_{k=0}^{N} \left[\underline{\hat{P}}_1^{\pm}(\zeta, r, u|k) \cos[k(\phi_v - \phi_s)] + \underline{\hat{P}}_2^{\pm}(\zeta, r, u|k) \sin[k(\phi_v - \phi_s)] \right] \right\} \\
&\qquad \times \left\{ \sum_{\ell=0}^{N} \left[\underline{\hat{S}}_1^{-}(\zeta, r|\ell) \cos(\ell\phi_s) + \underline{\hat{S}}_2^{-}(\zeta, r|\ell) \sin(\ell\phi_s) \right] \right\} \\
&+ \omega_o(\zeta) \underline{\hat{P}}_1^{\pm}(\zeta, M, u|0) \underline{\hat{S}}_1^{-}(\zeta, M|0) \\
&+ \omega_o(\zeta) \sum_{r=1}^{M-1} \sum_{s=1}^{2n} \left\{ \sum_{k=0}^{N} \left[\underline{\hat{P}}_1^{\mp}(\zeta, r, u|k) \cos[k(\phi_v - \phi_s)] + \underline{\hat{P}}_2^{\mp}(\zeta, r, u|k) \sin[k(\phi_v - \phi_s)] \right] \right\} \\
&\qquad \times \left\{ \sum_{\ell=0}^{N} \left[\underline{\hat{S}}_1^{+}(\zeta, r|\ell) \cos(\ell\phi_s) + \underline{\hat{S}}_2^{+}(\zeta, r|\ell) \sin(\ell\phi_s) \right] \right\} \\
&+ \omega_o(\zeta) \underline{\hat{P}}_1^{\mp}(\zeta, M, u|0) \underline{\hat{S}}_1^{+}(\zeta, M|0) \\
&+ \sum_{\ell=0}^{N} \left[\underline{\hat{\Sigma}}_1^{\mp}(\zeta, u|\ell) \cos(\ell\phi_v) + \underline{\hat{\Sigma}}_2^{\mp}(\zeta, u|\ell) \sin(\ell\phi_v) \right] \ . \qquad (4.6)
\end{aligned}
$$

Here the generic sums over r and s have been written to explicitly separate out the polar cap quads, $r = M$, which have no azimuthal dependence. For polar caps, the Fourier cosine amplitude for $k = 0$ is just the value of the physical variable, the cosine amplitudes are zero for $k > 0$, and the sine amplitudes are zero for all k values.

Equation (4.6) is more complicated than its scalar equivalent (L&W 8.35) because it is now necessary to have both cosine ($\underline{\hat{P}}_1^{\pm}$) and sine ($\underline{\hat{P}}_2^{\pm}$) amplitudes to represent $\underline{P}^{\pm}$, whereas the phase function $\tilde{\beta}$ in the SRTE needed only cosine amplitudes $\hat{\beta}$, so no subscript was needed on $\hat{\beta}$ as seen in L&W (8.35). The necessity of separating cosine and sign amplitudes carries through to the local and global transmission and reflection operators, as will be seen below.

The path function terms in Eq. (4.6) have quantities of the form

$$
\sum_{l=0}^{N} \sum_{k=0}^{N} f(k) g(\ell) \sum_{s=1}^{2N} \begin{Bmatrix} \cos[k(\phi_v - \phi_s)] \cos(\ell\phi_s) \\ \cos[k(\phi_v - \phi_s)] \sin(\ell\phi_s) \\ \sin[k(\phi_v - \phi_s)] \cos(\ell\phi_s) \\ \sin[k(\phi_v - \phi_s)] \sin(\ell\phi_s) \end{Bmatrix} ,
$$

where $f(k)$ is a phase matrix element for a given ζ, r, u, and $g(\ell)$ is a Stokes vector element for a given ζ, r. These four combinations of cosine and sine terms are reduced as in Eqs. (A.19) to

(A.22) to obtain [1]

$$\sum_{l=0}^{N} f(\ell)g(\ell) \begin{Bmatrix} \epsilon_\ell \cos(\ell\phi_v) \\ \gamma_\ell \sin(\ell\phi_v) \\ \epsilon_\ell \sin(\ell\phi_v) \\ -\gamma_\ell \cos(\ell\phi_v) \end{Bmatrix},$$

where ϵ_ℓ and γ_ℓ are defined by Eqs. (A.8) and (A.10):

$$\epsilon_\ell = N(1 + \delta_{2\ell} + \delta_{2\ell-2N}) = \begin{cases} 2N & \text{if } \ell = 0 \text{ or } \ell = N \\ N & \text{if } \ell = 1, 2, ..., N-1 \end{cases}$$

and

$$\gamma_\ell = N(1 - \delta_{2\ell} - \delta_{2\ell-2N}) = \begin{cases} 0 & \text{if } \ell = 0 \text{ or } \ell = N \\ N & \text{if } \ell = 1, 2, ..., N-1\,. \end{cases}$$

The Kronecker delta functions are defined by Eqs. (A.1) and (A.2).

Equation (4.6) therefore reduces to (L&W 8.36)

$$\begin{aligned}
&\sum_{\ell=0}^{N} \left\{ \mp\mu_u \frac{d}{d\zeta} \underline{\hat{S}}_1^{\mp}(\zeta, u|\ell) \cos(\ell\phi_v) \mp \mu_u \frac{d}{d\zeta} \underline{\hat{S}}_2^{\mp}(\zeta, u|\ell) \sin(\ell\phi_v) \right\} \\
&= -\sum_{\ell=0}^{N} \left\{ \underline{\hat{S}}_1^{\mp}(\zeta, u|\ell) \cos(\ell\phi_v) + \underline{\hat{S}}_2^{\mp}(\zeta, u|\ell) \sin(\ell\phi_v) \right\} \\
&+ \omega_o(\zeta) \sum_{\ell=0}^{N} \left\{ \sum_{r=1}^{M-1} \epsilon_\ell \, \underline{\hat{P}}_1^{\pm}(\zeta, r, u|\ell) \underline{\hat{S}}_1^{-}(\zeta, r|\ell) + \delta_\ell \, \underline{\hat{P}}_1^{\pm}(\zeta, M, u|\ell) \underline{\hat{S}}_1^{-}(\zeta, M|\ell) \right\} \cos(\ell\phi_v) \\
&- \omega_o(\zeta) \sum_{\ell=0}^{N} \left\{ \sum_{r=1}^{M-1} \gamma_\ell \, \underline{\hat{P}}_2^{\pm}(\zeta, r, u|\ell) \underline{\hat{S}}_2^{-}(\zeta, r|\ell) \right\} \cos(\ell\phi_v) \\
&+ \omega_o(\zeta) \sum_{\ell=0}^{N} \left\{ \sum_{r=1}^{M-1} \gamma_\ell \, \underline{\hat{P}}_1^{\pm}(\zeta, r, u|\ell) \underline{\hat{S}}_2^{-}(\zeta, r|\ell) + \sum_{r=1}^{M-1} \epsilon_\ell \, \underline{\hat{P}}_2^{\pm}(\zeta, r, u|\ell) \underline{\hat{S}}_1^{-}(\zeta, r|\ell) \right\} \sin(\ell\phi_v) \\
&+ \omega_o(\zeta) \sum_{\ell=0}^{N} \left\{ \sum_{r=1}^{M-1} \epsilon_\ell \, \underline{\hat{P}}_1^{\mp}(\zeta, r, u|\ell) \underline{\hat{S}}_1^{+}(\zeta, r|\ell) + \delta_\ell \, \underline{\hat{P}}_1^{\pm}(\zeta, M, u|\ell) \underline{\hat{S}}_1^{+}(\zeta, M|\ell) \right\} \cos(\ell\phi_v) \\
&- \omega_o(\zeta) \sum_{\ell=0}^{N} \left\{ \sum_{r=1}^{M-1} \gamma_\ell \, \underline{\hat{P}}_2^{\mp}(\zeta, r, u|\ell) \underline{\hat{S}}_2^{+}(\zeta, r|\ell) \right\} \cos(\ell\phi_v) \\
&+ \omega_o(\zeta) \sum_{\ell=0}^{N} \left\{ \sum_{r=1}^{M-1} \gamma_\ell \, \underline{\hat{P}}_1^{\mp}(\zeta, r, u|\ell) \underline{\hat{S}}_2^{+}(\zeta, r|\ell) + \sum_{r=1}^{M-1} \epsilon_\ell \, \underline{\hat{P}}_2^{\mp}(\zeta, r, u|\ell) \underline{\hat{S}}_1^{+}(\zeta, r|\ell) \right\} \sin(\ell\phi_v) \\
&+ \sum_{\ell=0}^{N} \left\{ \underline{\hat{\Sigma}}_1^{\mp}(\zeta, u|\ell) \cos(\ell\phi_v) + \underline{\hat{\Sigma}}_2^{\mp}(\zeta, u|\ell) \sin(\ell\phi_v) \right\} \,. \qquad (4.7)
\end{aligned}$$

[1] The minus sign in $-\gamma_\ell \cos(\ell\phi_v)$ results from the reduction of the $\sin[k(\phi_v - \phi_s)] \sin(\ell\phi_s)$ term and the minus sign in $\sin(a - b) = \sin a \cos b - \cos a \sin b$, so there's no escaping it. As will be seen, this sign carries through the rest of the derivation and causes a certain loss of symmetry in many equations, compared to the scalar equations in *Light and Water*.

The polar cap terms, which involve only $\ell = 0$ cosine amplitudes, have been incorporated into the sums over ℓ by use of δ_ℓ. Note that the minus sign in resulting from the reduction of the sums over $\sin[k(\phi_v - \phi_s)]\sin(\ell\phi_s)$ gives a minus sign in all terms involving $\underline{\hat{P}}_2^{\pm}(\zeta, r, u|\ell)\underline{\hat{S}}_2^{\pm}(\zeta, r|\ell)$. This sign carries through the rest of the derivation and introduces a sign asymmetry in the cosine and sine equations, which does not occur in the corresponding scalar equations (which do not require $\underline{\hat{P}}_2^{\pm}(\zeta, r, u|\ell)$ terms in the Fourier expansion of the phase function). The same sign appears, for example, in Eq. (16) of Hovenier (1971) in a decomposition of the continuous-variable (undiscretized) VRTE.

The linear independence of $\cos(\ell\phi_v)$ and $\sin(\ell\phi_v)$ for different ℓ values means that this last equation must separately hold true for each value of $\ell = 0, ..., N$ for the $\underline{\hat{S}}_1$ amplitudes, and for $\ell = 1, ..., N-1$ for the $\underline{\hat{S}}_2$ amplitudes. Collecting together the coefficients of $\cos(\ell\phi_v)$ then gives

$$
\begin{aligned}
&\mp\mu_u \frac{d}{d\zeta}\underline{\hat{S}}_1^{\mp}(\zeta, u|\ell) = -\underline{\hat{S}}_1^{\mp}(\zeta, u|\ell) \\
&+ \sum_{r=1}^{M-1} \omega_o(\zeta)\,\epsilon_\ell\,\underline{\hat{P}}_1^{\pm}(\zeta, r, u|\ell)\underline{\hat{S}}_1^{-}(\zeta, r|\ell) + \omega_o(\zeta)\,\delta_\ell\,\underline{\hat{P}}_1^{\pm}(\zeta, M, u|\ell)\underline{\hat{S}}_1^{-}(\zeta, M|\ell) \\
&- \sum_{r=1}^{M-1} \omega_o(\zeta)\,\gamma_\ell\,\underline{\hat{P}}_2^{\pm}(\zeta, r, u|\ell)\underline{\hat{S}}_2^{-}(\zeta, r|\ell) \\
&+ \sum_{r=1}^{M-1} \omega_o(\zeta)\,\epsilon_\ell\,\underline{\hat{P}}_1^{\mp}(\zeta, r, u|\ell)\underline{\hat{S}}_1^{+}(\zeta, r|\ell) + \omega_o(\zeta)\,\delta_\ell\,\underline{\hat{P}}_1^{\mp}(\zeta, M, u|\ell)\underline{\hat{S}}_1^{+}(\zeta, M|\ell) \\
&- \sum_{r=1}^{M-1} \omega_o(\zeta)\,\gamma_\ell\,\underline{\hat{P}}_2^{\mp}(\zeta, r, u|\ell)\underline{\hat{S}}_2^{+}(\zeta, r|\ell) \\
&+ \underline{\hat{\Sigma}}_1^{\mp}(\zeta, u|\ell)\ . \qquad (4.8)
\end{aligned}
$$

The equation obtained from the coefficients of the $\sin(\ell\phi_v)$ terms is

$$
\begin{aligned}
&\mp\mu_u \frac{d}{d\zeta}\underline{\hat{S}}_2^{\mp}(\zeta, u|\ell) = -\underline{\hat{S}}_2^{\mp}(\zeta, u|\ell) \\
&+ \sum_{r=1}^{M-1} \omega_o(\zeta)\,\gamma_\ell\,\underline{\hat{P}}_1^{\pm}(\zeta, r, u|\ell)\underline{\hat{S}}_2^{-}(\zeta, r|\ell) + \sum_{r=1}^{M-1} \omega_o(\zeta)\,\epsilon_\ell\,\underline{\hat{P}}_2^{\pm}(\zeta, r, u|\ell)\underline{\hat{S}}_1^{-}(\zeta, r|\ell) \\
&+ \sum_{r=1}^{M-1} \omega_o(\zeta)\,\gamma_\ell\,\underline{\hat{P}}_1^{\mp}(\zeta, r, u|\ell)\underline{\hat{S}}_2^{+}(\zeta, r|\ell) + \sum_{r=1}^{M-1} \omega_o(\zeta)\,\epsilon_\ell\,\underline{\hat{P}}_2^{\mp}(\zeta, r, u|\ell)\underline{\hat{S}}_1^{+}(\zeta, r|\ell) \\
&+ \underline{\hat{\Sigma}}_2^{\mp}(\zeta, u|\ell)\ . \qquad (4.9)
\end{aligned}
$$

Since $\underline{\hat{S}}_2, \underline{\hat{P}}_2$, and $\underline{\hat{\Sigma}}_2$ are all 0 when $\ell = 0$ or $\ell = N$, we can for notational convenience regard this last equation as holding for $\ell = 0$ and N as well. Thus Eqs. (4.8) and (4.8) can both be regarded as holding for $u = 1, ..., M-1$ and $\ell = 0, ..., N$.

Note in these equations that $\underline{\hat{P}}_1$ scatters $\underline{\hat{S}}_1$ into $\underline{\hat{S}}_1$ and $\underline{\hat{S}}_2$ into $\underline{\hat{S}}_2$, whereas $\underline{\hat{P}}_2$ scatters $\underline{\hat{S}}_1$ into $\underline{\hat{S}}_2$ and $\underline{\hat{S}}_2$ into $\underline{\hat{S}}_1$. That is, $\underline{\hat{P}}_2$ converts Stokes vector cosine amplitudes into sine amplitudes, and vice versa.

The equation for the polar cap $u = M$ involves only the cosine term for $\ell = 0$ since there is no

ϕ dependency. The vector equations are therefore have the same form as the scalar (L&W 8.39):

$$\begin{aligned}
&\mp\mu_u \frac{d}{d\zeta}\underline{\hat{S}}_1^{\mp}(\zeta, M|0) = -\underline{\hat{S}}_1^{\mp}(\zeta, M|0) \\
&+ \sum_{r=1}^{M-1} \omega_o(\zeta)\,\epsilon_0\,\underline{\hat{P}}_1^{\pm}(\zeta, r, M|0)\underline{\hat{S}}_1^{-}(\zeta, r|0) + \omega_o(\zeta)\,\underline{\hat{P}}_1^{\pm}(\zeta, M, M|0)\underline{\hat{S}}_1^{-}(\zeta, M|0) \\
&+ \sum_{r=1}^{M-1} \omega_o(\zeta)\,\epsilon_0\,\underline{\hat{P}}_1^{\mp}(\zeta, r, M|0)\underline{\hat{S}}_1^{-}(\zeta, r|0) + \omega_o(\zeta)\,\underline{\hat{P}}_1^{\mp}(\zeta, M, M|0)\underline{\hat{S}}_1^{+}(\zeta, M|0) \\
&+ \underline{\hat{\Sigma}}_1^{\mp}(\zeta, M|0)\ . \qquad (4.10)
\end{aligned}$$

Taking the lower signs in Eq. (4.8) gives the equation for the downwelling radiance cosine amplitudes. The "attenuation" term $-\underline{\hat{S}}_1^{+}(\zeta, u|\ell)$ term can be incorporated into the sum over r by use of a Kronecker delta function. The result is (L&W 8.40)

$$\begin{aligned}
\frac{d}{d\zeta}\underline{\hat{S}}_1^{+}(\zeta, u|\ell) = &\sum_{r=1}^{M-1}\left[\frac{\omega_o(\zeta)\epsilon_\ell}{\mu_u}\underline{\hat{P}}_1^{-}(\zeta, r, u|\ell)\right]\underline{\hat{S}}_1^{-}(\zeta, r|\ell) + \left[\frac{\omega_o(\zeta)\delta_\ell}{\mu_u}\underline{\hat{P}}_1^{-}(\zeta, M, u|\ell)\right]\underline{\hat{S}}_1^{-}(\zeta, M|\ell) \\
&- \sum_{r=1}^{M-1}\left[\frac{\omega_o(\zeta)\gamma_\ell}{\mu_u}\underline{\hat{P}}_2^{-}(\zeta, r, u|\ell)\right]\underline{\hat{S}}_2^{-}(\zeta, r|\ell) \\
&+ \sum_{r=1}^{M-1}\left[\frac{\omega_o(\zeta)\epsilon_\ell}{\mu_u}\underline{\hat{P}}_1^{+}(\zeta, r, u|\ell) - \frac{\delta_{r-u}}{\mu_u}\underline{I}_4\right]\underline{\hat{S}}_1^{+}(\zeta, r|\ell) + \left[\frac{\omega_o(\zeta)\delta_\ell}{\mu_u}\underline{\hat{P}}_1^{+}(\zeta, M, u|\ell)\right]\underline{\hat{S}}_1^{+}(\zeta, M|\ell) \\
&- \sum_{r=1}^{M-1}\left[\frac{\omega_o(\zeta)\gamma_\ell}{\mu_u}\underline{\hat{P}}_2^{+}(\zeta, r, u|\ell)\right]\underline{\hat{S}}_2^{+}(\zeta, r|\ell) + \frac{1}{\mu_u}\underline{\hat{\Sigma}}_1^{+}(\zeta, u|\ell)\,. \qquad (4.11)
\end{aligned}$$

where $\underline{I}_4$ is the 4×4 identity matrix. The corresponding equation for the upwelling radiance cosine amplitudes is (L&W 8.41)

$$\begin{aligned}
-\frac{d}{d\zeta}\underline{\hat{S}}_1^{-}(\zeta, u|\ell) = &\sum_{r=1}^{M-1}\left[\frac{\omega_o(\zeta)\epsilon_\ell}{\mu_u}\underline{\hat{P}}_1^{+}(\zeta, r, u|\ell) - \frac{\delta_{r-u}}{\mu_u}\underline{I}_4\right]\underline{\hat{S}}_1^{-}(\zeta, r|\ell) + \left[\frac{\omega_o(\zeta)\delta_\ell}{\mu_u}\underline{\hat{P}}_1^{+}(\zeta, M, u|\ell)\right]\underline{\hat{S}}_1^{-}(\zeta, M|\ell) \\
&- \sum_{r=1}^{M-1}\left[\frac{\omega_o(\zeta)\gamma_\ell}{\mu_u}\underline{\hat{P}}_2^{+}(\zeta, r, u|\ell)\right]\underline{\hat{S}}_2^{-}(\zeta, r|\ell) \\
&+ \sum_{r=1}^{M-1}\left[\frac{\omega_o(\zeta)\epsilon_\ell}{\mu_u}\underline{\hat{P}}_1^{-}(\zeta, r, u|\ell)\right]\underline{\hat{S}}_1^{+}(\zeta, r|\ell) + \left[\frac{\omega_o(\zeta)\delta_\ell}{\mu_u}\underline{\hat{P}}_1^{-}(\zeta, M, u|\ell)\right]\underline{\hat{S}}_1^{+}(\zeta, M|\ell) \\
&- \sum_{r=1}^{M-1}\left[\frac{\omega_o(\zeta)\gamma_\ell}{\mu_u}\underline{\hat{P}}_2^{-}(\zeta, r, u|\ell)\right]\underline{\hat{S}}_2^{+}(\zeta, r|\ell) + \frac{1}{\mu_u}\underline{\hat{\Sigma}}_1^{-}(\zeta, u|\ell)\,. \qquad (4.12)
\end{aligned}$$

For these cosine amplitudes, $u = 1, ..., M-1$ and $\ell = 0, ..., N$.

The corresponding equation for the downwelling sine amplitudes is

$$\begin{aligned}
\frac{d}{d\zeta}\underline{\hat{S}}_2^+(\zeta,u|\ell) = & \sum_{r=1}^{M-1}\left[\frac{\omega_o(\zeta)\gamma_\ell}{\mu_u}\underline{\hat{P}}_1^-(\zeta,r,u|\ell)\right]\underline{\hat{S}}_2^-(\zeta,r|\ell) \\
& + \sum_{r=1}^{M-1}\left[\frac{\omega_o(\zeta)\epsilon_\ell}{\mu_u}\underline{\hat{P}}_2^-(\zeta,r,u|\ell)\right]\underline{\hat{S}}_1^-(\zeta,r|\ell) \\
& + \sum_{r=1}^{M-1}\left[\frac{\omega_o(\zeta)\gamma_\ell}{\mu_u}\underline{\hat{P}}_1^+(\zeta,r,u|\ell) - \frac{\delta_{r-u}}{\mu_u}\underline{I}_4\right]\underline{\hat{S}}_2^+(\zeta,r|\ell) \\
& + \sum_{r=1}^{M-1}\left[\frac{\omega_o(\zeta)\epsilon_\ell}{\mu_u}\underline{\hat{P}}_2^+(\zeta,r,u|\ell)\right]\underline{\hat{S}}_1^+(\zeta,r|\ell) + \frac{1}{\mu_u}\underline{\hat{\Sigma}}_2^+(\zeta,u|\ell)\,, \qquad (4.13)
\end{aligned}$$

and the equation for the upwelling sine amplitudes is

$$\begin{aligned}
-\frac{d}{d\zeta}\underline{\hat{S}}_2^-(\zeta,u|\ell) = & \sum_{r=1}^{M-1}\left[\frac{\omega_o(\zeta)\gamma_\ell}{\mu_u}\underline{\hat{P}}_1^+(\zeta,r,u|\ell) - \frac{\delta_{r-u}}{\mu_u}\underline{I}_4\right]\underline{\hat{S}}_2^-(\zeta,r|\ell) \\
& + \sum_{r=1}^{M-1}\left[\frac{\omega_o(\zeta)\epsilon_\ell}{\mu_u}\underline{\hat{P}}_2^+(\zeta,r,u|\ell)\right]\underline{\hat{S}}_1^-(\zeta,r|\ell) \\
& + \sum_{r=1}^{M-1}\left[\frac{\omega_o(\zeta)\gamma_\ell}{\mu_u}\underline{\hat{P}}_1^-(\zeta,r,u|\ell)\right]\underline{\hat{S}}_2^+(\zeta,r|\ell) \\
& + \sum_{r=1}^{M-1}\left[\frac{\omega_o(\zeta)\epsilon_\ell}{\mu_u}\underline{\hat{P}}_2^-(\zeta,r,u|\ell)\right]\underline{\hat{S}}_1^+(\zeta,r|\ell) + \frac{1}{\mu_u}\underline{\hat{\Sigma}}_2^-(\zeta,u|\ell)\,. \qquad (4.14)
\end{aligned}$$

For the sine amplitude equations, $u = 1, ..., M-1$ and $\ell = 1, ..., N-1$. Because $\epsilon_\ell = \gamma_\ell = N$ for $\ell = 1, ..., N-1$, the sine equations can be written using ϵ_ℓ.

4.3.1 Local Reflectance and Transmittance Matrices

The quantities in brackets in Eqs. (4.11) to (4.14) are the *local reflectance* and *local transmittance* functions for the radiance amplitudes. There are several things to note about these quantities:

- These quantities depend only the IOPs at depth ζ (and wavelength λ) and on the choice of quad partioning. They are thus "local" IOPs at a particular depth.
- The same quantities appear in both the downward and upward equations.
- The same quantities appear in both cosine and sine equations when $\ell = 1, ..., N-1$, in which case $\epsilon_\ell = \gamma_\ell$
- The terms involving $\underline{\hat{P}}_{1,2}^+$ "transmit" downwelling radiance into downwelling radiance, and upwelling into upwelling.
- The terms involving $\underline{\hat{P}}_{1,2}^-$ "reflect" downwelling radiance into upwelling radiance, and upwelling into downwelling.

It will prove convenient to cast the preceding equations into matrix form. Recall that the phase matrix amplitudes are 4×4 for a given (r, u) value. Next define $M \times M$ *composite* matrices $\underline{\underline{\hat{\rho}}}^{\pm}_{1,2}$ and $\underline{\underline{\hat{\tau}}}^{\pm}_{1,2}$ whose (row, column) elements (u, r) are obtained from the 4×4 $\underline{\hat{P}}^{\pm}_{1,2}(\zeta, r, u|\ell)$ matrices in the brackets. That is, $\underline{\underline{\hat{\rho}}}$ and $\underline{\underline{\hat{\tau}}}$ are $M \times M$ matrices, each of whose elements is a 4×4 matrix, so that the total size of $\underline{\underline{\hat{\rho}}}$ and $\underline{\underline{\hat{\tau}}}$ is $4M \times 4M$ when written out in full as real numbers. The double underline is a reminder that $\underline{\underline{\hat{\rho}}}$ and $\underline{\underline{\hat{\tau}}}$ are matrices of matrices. These matrices are defined as follows: Quad-to-quad case, $r, u = 1, ..., M-1$:

$$[\,\underline{\underline{\hat{\tau}}}_1(\zeta|\ell)]_{u,r} \equiv \frac{\omega_o(\zeta)\epsilon_\ell}{\mu_u}\underline{\hat{P}}^+_1(\zeta, r, u|\ell) - \frac{\delta_{r-u}}{\mu_u}\underline{I}_4 \qquad \text{for } \ell = 0, ..., \mathrm{N} \tag{4.15a}$$

$$[\,\underline{\underline{\hat{\tau}}}_2(\zeta|\ell)]_{u,r} \equiv \frac{\omega_o(\zeta)\gamma_\ell}{\mu_u}\underline{\hat{P}}^+_2(\zeta, r, u|\ell) \qquad \text{for } \ell = 1, ..., \mathrm{N}-1 \tag{4.15b}$$

$$[\,\underline{\underline{\hat{\tau}}}_2(\zeta|\ell)]_{u,r} \equiv \underline{0}_4 \qquad \text{for } \ell = 0 \text{ or N} \tag{4.15c}$$

$$[\,\underline{\underline{\hat{\rho}}}_1(\zeta|\ell)]_{u,r} \equiv \frac{\omega_o(\zeta)\epsilon_\ell}{\mu_u}\underline{\hat{P}}^-_1(\zeta, r, u|\ell) \qquad \text{for } \ell = 0, ..., \mathrm{N} \tag{4.15d}$$

$$[\,\underline{\underline{\hat{\rho}}}_2(\zeta|\ell)]_{u,r} \equiv \frac{\omega_o(\zeta)\gamma_\ell}{\mu_u}\underline{\hat{P}}^-_2(\zeta, r, u|\ell) \qquad \text{for } \ell = 1, ..., \mathrm{N}-1 \tag{4.15e}$$

$$[\,\underline{\underline{\hat{\rho}}}_2(\zeta|\ell)]_{u,r} \equiv \underline{0}_4 \qquad \text{for } \ell = 0 \text{ or N} \tag{4.15f}$$

Polar cap-to-quad case, $r = M, u = 1, ..., M-1$:

$$[\,\underline{\underline{\hat{\tau}}}_1(\zeta|0)]_{u,M} \equiv \frac{\omega_o(\zeta)}{\mu_u}\underline{\hat{P}}^+_1(\zeta, M, u|0) \qquad \text{for } \ell = 0 \tag{4.16a}$$

$$[\,\underline{\underline{\hat{\tau}}}_1(\zeta|\ell)]_{u,M} \equiv \underline{0}_4 \qquad \text{for } \ell = 1, ..., \mathrm{N} \tag{4.16b}$$

$$[\,\underline{\underline{\hat{\tau}}}_2(\zeta|\ell)]_{u,M} \equiv \underline{0}_4 \qquad \text{for} \qquad \text{for } \ell = 0, ..., \mathrm{N} \tag{4.16c}$$

$$[\,\underline{\underline{\hat{\rho}}}_1(\zeta|0)]_{u,M} \equiv \frac{\omega_o(\zeta)}{\mu_u}\underline{\hat{P}}^-_1(\zeta, M, u|0) \qquad \text{for } \ell = 0 \tag{4.16d}$$

$$[\,\underline{\underline{\hat{\rho}}}_1(\zeta|\ell)]_{u,M} \equiv \underline{0}_4 \qquad \text{for } \ell = 1, ..., \mathrm{N} \tag{4.16e}$$

$$[\,\underline{\underline{\hat{\rho}}}_2(\zeta|\ell)]_{u,M} \equiv \underline{0}_4 \qquad \text{for} \qquad \text{for } \ell = 0, ..., \mathrm{N} \tag{4.16f}$$

Quad-to-polar cap case, $r = 1, ..., M-1, u = M$:

$$[\,\underline{\underline{\hat{\tau}}}_1(\zeta|0)]_{M,r} \equiv \frac{\omega_o(\zeta)\epsilon_0}{\mu_u}\underline{\hat{P}}^+_1(\zeta, r, M|0) \tag{4.17a}$$

$$[\,\underline{\underline{\hat{\tau}}}_1(\zeta|\ell)]_{M,r} \equiv \underline{0}_4 \qquad \text{for } \ell = 1, ..., \mathrm{N} \tag{4.17b}$$

$$[\,\underline{\underline{\hat{\tau}}}_2(\zeta|\ell)]_{M,r} \equiv \underline{0}_4 \qquad \text{for } \ell = 0, ..., \mathrm{N} \tag{4.17c}$$

$$[\,\underline{\underline{\hat{\rho}}}_1(\zeta|0)]_{M,r} \equiv \frac{\omega_o(\zeta)\epsilon_0}{\mu_u}\underline{\hat{P}}^-_1(\zeta, r, M|0) \tag{4.17d}$$

$$[\,\underline{\underline{\hat{\rho}}}_1(\zeta|\ell)]_{M,r} \equiv \underline{0}_4 \qquad \text{for } \ell = 1, ..., \mathrm{N} \tag{4.17e}$$

$$[\,\underline{\underline{\hat{\rho}}}_2(\zeta|\ell)]_{M,r} \equiv \underline{0}_4 \qquad \text{for } \ell = 0, ..., \mathrm{N} \tag{4.17f}$$

Polar cap-to-polar cap case, $r = M, u = M$:

$$[\underline{\hat{\tau}}_1(\zeta|0)]_{M,M} \equiv \frac{\omega_o(\zeta)\epsilon_0}{\mu_u}\underline{\hat{P}}_1^+(\zeta, M, M|0) - \frac{1}{\mu_M}\underline{I}_4 \tag{4.18a}$$

$$[\underline{\hat{\tau}}_1(\zeta|\ell)]_{M,M} \equiv \underline{0}_4 \qquad \text{for } \ell = 1, ..., \mathrm{N} \tag{4.18b}$$

$$[\underline{\hat{\tau}}_2(\zeta|\ell)]_{M,M} \equiv \underline{0}_4 \qquad \text{for } \ell = 0, ..., \mathrm{N} \tag{4.18c}$$

$$[\underline{\hat{\rho}}_1(\zeta|0)]_{M,M} \equiv \frac{\omega_o(\zeta)\epsilon_0}{\mu_u}\underline{\hat{P}}_1^-(\zeta, M, M|0) \tag{4.18d}$$

$$[\underline{\hat{\rho}}_1(\zeta|\ell)]_{M,M} \equiv \underline{0}_4 \qquad \text{for } \ell = 1, ..., \mathrm{N} \tag{4.18e}$$

$$[\underline{\hat{\rho}}_2(\zeta|\ell)]_{M,M} \equiv \underline{0}_4 \qquad \text{for } \ell = 0, ..., \mathrm{N} \tag{4.18f}$$

Note that the elements of the M^{th} row and M^{th} column of $\underline{\hat{\rho}}_1$ and $\underline{\hat{\tau}}_1$ are $\underline{0}_4$ except when $\ell = 0$; $\underline{0}_4$ is the 4×4 matrix of zeros.

With these definitions, Eqs. (4.11) and (4.12) can be written

$$\begin{aligned}\frac{d}{d\zeta}\underline{\hat{S}}_1^+(\zeta, u|\ell) = &\sum_{r=1}^{M}[\underline{\hat{\rho}}_1(\zeta|\ell)]_{u,r}\,\underline{\hat{S}}_1^-(\zeta, r|\ell) + \sum_{r=1}^{M}[\underline{\hat{\tau}}_1(\zeta|\ell)]_{u,r}\,\underline{\hat{S}}_1^+(\zeta, r|\ell) \\ &- \sum_{r=1}^{M}[\underline{\hat{\rho}}_2(\zeta|\ell)]_{u,r}\,\underline{\hat{S}}_2^-(\zeta, r|\ell) - \sum_{r=1}^{M}[\underline{\hat{\tau}}_2(\zeta|\ell)]_{u,r}\,\underline{\hat{S}}_2^+(\zeta, r|\ell) + \frac{1}{\mu_u}\underline{\hat{\Sigma}}_1^+(\zeta, u|\ell)\end{aligned} \tag{4.19}$$

and

$$\begin{aligned}-\frac{d}{d\zeta}\underline{\hat{S}}_1^-(\zeta, u|\ell) = &\sum_{r=1}^{M}[\underline{\hat{\rho}}_1(\zeta|\ell)]_{u,r}\,\underline{\hat{S}}_1^+(\zeta, r|\ell) + \sum_{r=1}^{M}[\underline{\hat{\tau}}_1(\zeta|\ell)]_{u,r}\,\underline{\hat{S}}_1^-(\zeta, r|\ell) \\ &- \sum_{r=1}^{M}[\underline{\hat{\rho}}_2(\zeta|\ell)]_{u,r}\,\underline{\hat{S}}_2^+(\zeta, r|\ell) - \sum_{r=1}^{M}[\underline{\hat{\tau}}_2(\zeta|\ell)]_{u,r}\,\underline{\hat{S}}_2^-(\zeta, r|\ell) + \frac{1}{\mu_u}\underline{\hat{\Sigma}}_1^-(\zeta, u|\ell)\end{aligned} \tag{4.20}$$

These equations are valid for $\ell = 0, ..., N$ and $u = 1, ..., M-1$. The polar caps Eq. (4.10) results in a similar pair of equations involving $\underline{\hat{\rho}}_1(\zeta, 0)_{M,r}$ and $\underline{\hat{\tau}}_1(\zeta, 0)_{M,r}$.

The corresponding equations for the sine amplitudes are

$$\begin{aligned}\frac{d}{d\zeta}\underline{\hat{S}}_2^+(\zeta, u|\ell) = &\sum_{r=1}^{M}[\underline{\hat{\rho}}_1(\zeta|\ell)]_{u,r}\,\underline{\hat{S}}_2^-(\zeta, r|\ell) + \sum_{r=1}^{M}[\underline{\hat{\tau}}_1(\zeta|\ell)]_{u,r}\,\underline{\hat{S}}_2^+(\zeta, r|\ell) \\ &+ \sum_{r=1}^{M}[\underline{\hat{\rho}}_2(\zeta|\ell)]_{u,r}\,\underline{\hat{S}}_1^-(\zeta, r|\ell) + \sum_{r=1}^{M}[\underline{\hat{\tau}}_2(\zeta|\ell)]_{u,r}\,\underline{\hat{S}}_1^+(\zeta, r|\ell) + \frac{1}{\mu_u}\underline{\hat{\Sigma}}_2^+(\zeta, u|\ell)\end{aligned} \tag{4.21}$$

and

$$\begin{aligned}-\frac{d}{d\zeta}\underline{\hat{S}}_2^-(\zeta, u|\ell) = &\sum_{r=1}^{M}[\underline{\hat{\rho}}_1(\zeta|\ell)]_{u,r}\,\underline{\hat{S}}_2^+(\zeta, r|\ell) + \sum_{r=1}^{M}[\underline{\hat{\tau}}_1(\zeta|\ell)]_{u,r}\,\underline{\hat{S}}_2^-(\zeta, r|\ell) \\ &+ \sum_{r=1}^{M}[\underline{\hat{\rho}}_2(\zeta|\ell)]_{u,r}\,\underline{\hat{S}}_1^+(\zeta, r|\ell) + \sum_{r=1}^{M}[\underline{\hat{\tau}}_2(\zeta|\ell)]_{u,r}\,\underline{\hat{S}}_1^-(\zeta, r|\ell) + \frac{1}{\mu_u}\underline{\hat{\Sigma}}_2^-(\zeta, u|\ell)\end{aligned} \tag{4.22}$$

These equations are valid for $\ell = 1, ..., N-1$ and $u = 1, ..., M-1$. The sine amplitudes are zero for the polar caps $u = M$.

It can now be seen that the (u, r) = (row, column) order in the definitions of $[\underline{\underline{\hat{\rho}}}_{1,2}(\zeta|\ell)]_{u,r}$ and $[\underline{\underline{\hat{\tau}}}_{1,2}(\zeta|\ell)]_{u,r}$ was chosen to facilitate converting the summations over r in Eqs. (4.11) to (4.14) to matrix multiplications. Continuing in this direction, now "stack up" the Stokes vectors to define $4M \times 1$ composite column vectors:

$$\underline{\hat{S}}^{\pm}_{1,2}(\zeta|\ell) = \begin{bmatrix} \underline{\hat{S}}^{\pm}_{1,2}(\zeta, 1|\ell) \\ \underline{\hat{S}}^{\pm}_{1,2}(\zeta, 2|\ell) \\ \vdots \\ \underline{\hat{S}}^{\pm}_{1,2}(\zeta, M-1|\ell) \\ \underline{\hat{S}}^{\pm}_{1,2}(\zeta, M|\ell) \end{bmatrix}. \tag{4.23}$$

Similarly, arrange the 4×4 elements $[\underline{\underline{\hat{\rho}}}_{1,2}(\zeta|\ell)]_{u,r}$ of the composite $\underline{\underline{\hat{\rho}}}_{1,2}$ matrices into $4M \times 4M$ matrices

$$\underline{\underline{\hat{\rho}}}_{1,2}(\zeta|\ell) = \begin{bmatrix} [\underline{\underline{\hat{\rho}}}_{1,2}(\zeta|\ell)]_{1,1} & [\underline{\underline{\hat{\rho}}}_{1,2}(\zeta|\ell)]_{1,2} & \cdots & [\underline{\underline{\hat{\rho}}}_{1,2}(\zeta|\ell)]_{1,M} \\ [\underline{\underline{\hat{\rho}}}_{1,2}(\zeta|\ell)]_{2,1} & [\underline{\underline{\hat{\rho}}}_{1,2}(\zeta|\ell)]_{2,2} & \cdots & [\underline{\underline{\hat{\rho}}}_{1,2}(\zeta|\ell)]_{2,M} \\ \vdots & \vdots & \ddots & \vdots \\ [\underline{\underline{\hat{\rho}}}_{1,2}(\zeta|\ell)]_{M,1} & [\underline{\underline{\hat{\rho}}}_{1,2}(\zeta|\ell)]_{M,2} & \cdots & [\underline{\underline{\hat{\rho}}}_{1,2}(\zeta|\ell)]_{M,M} \end{bmatrix}. \tag{4.24}$$

with a similar equation for $\underline{\underline{\hat{\tau}}}_{1,2}(\zeta|\ell)$.

4.3.2 Matrix Form of the Local Interaction Equations

Definitions (4.23) and (4.24) allow Eqs. (4.19) to (4.22) to be written as matrix equations (L&W 8.43 and 8.44):

$$\frac{d}{d\zeta}\underline{\hat{S}}^{+}_{1}(\zeta|\ell) = \underline{\underline{\hat{\rho}}}_{1}(\zeta|\ell)\underline{\hat{S}}^{-}_{1}(\zeta|\ell) + \underline{\underline{\hat{\tau}}}_{1}(\zeta|\ell)\underline{\hat{S}}^{+}_{1}(\zeta|\ell) - \underline{\underline{\hat{\rho}}}_{2}(\zeta|\ell)\underline{\hat{S}}^{-}_{2}(\zeta|\ell) - \underline{\underline{\hat{\tau}}}_{2}(\zeta|\ell)\underline{\hat{S}}^{+}_{2}(\zeta|\ell) + \underline{\hat{\Sigma}}^{+}_{1}(\zeta|\ell) \tag{4.25}$$

$$-\frac{d}{d\zeta}\underline{\hat{S}}^{-}_{1}(\zeta|\ell) = \underline{\underline{\hat{\rho}}}_{1}(\zeta|\ell)\underline{\hat{S}}^{+}_{1}(\zeta|\ell) + \underline{\underline{\hat{\tau}}}_{1}(\zeta|\ell)\underline{\hat{S}}^{-}_{1}(\zeta|\ell) - \underline{\underline{\hat{\rho}}}_{2}(\zeta|\ell)\underline{\hat{S}}^{+}_{2}(\zeta|\ell) - \underline{\underline{\hat{\tau}}}_{2}(\zeta|\ell)\underline{\hat{S}}^{-}_{2}(\zeta|\ell) + \underline{\hat{\Sigma}}^{-}_{1}(\zeta|\ell) \tag{4.26}$$

$$\frac{d}{d\zeta}\underline{\hat{S}}^{+}_{2}(\zeta|\ell) = \underline{\underline{\hat{\rho}}}_{1}(\zeta|\ell)\underline{\hat{S}}^{-}_{2}(\zeta|\ell) + \underline{\underline{\hat{\tau}}}_{1}(\zeta|\ell)\underline{\hat{S}}^{+}_{2}(\zeta|\ell) + \underline{\underline{\hat{\rho}}}_{2}(\zeta|\ell)\underline{\hat{S}}^{-}_{1}(\zeta|\ell) + \underline{\underline{\hat{\tau}}}_{2}(\zeta|\ell)\underline{\hat{S}}^{+}_{1}(\zeta|\ell) + \underline{\hat{\Sigma}}^{+}_{2}(\zeta|\ell) \tag{4.27}$$

$$-\frac{d}{d\zeta}\underline{\hat{S}}^{-}_{2}(\zeta|\ell) = \underline{\underline{\hat{\rho}}}_{1}(\zeta|\ell)\underline{\hat{S}}^{+}_{2}(\zeta|\ell) + \underline{\underline{\hat{\tau}}}_{1}(\zeta|\ell)\underline{\hat{S}}^{-}_{2}(\zeta|\ell) + \underline{\underline{\hat{\rho}}}_{2}(\zeta|\ell)\underline{\hat{S}}^{+}_{1}(\zeta|\ell) + \underline{\underline{\hat{\tau}}}_{2}(\zeta|\ell)\underline{\hat{S}}^{-}_{1}(\zeta|\ell) + \underline{\hat{\Sigma}}^{-}_{2}(\zeta|\ell) \tag{4.28}$$

These equations are the composite matrix form of the *local interaction equations* for the radiance amplitudes. The minus signs on the $\underline{\underline{\hat{\rho}}}_{2}\underline{\hat{S}}^{\pm}_{2}$ and $\underline{\underline{\hat{\tau}}}_{2}\underline{\hat{S}}^{\pm}_{2}$ terms in Eqs. (4.25) and (4.26) trace back

to the minus sign in Eq. (A.22)[2]. These equations hold for each value of $\ell = 0, 1, ..., N$, although some terms may be zero for special cases, e.g., sine terms are zero when $\ell = 0$ or N.

To summarize, the development so far has

1. Written the discretized VRTE as upward and downward sets of equations,
2. Fourier decomposed the upward and downward equations,
3. Written the resulting equations in matrix form.

The local interaction equations are therefore just a rewritten version of the VRTE.

4.4 Global Interaction Equations

The local interaction equations show how radiance is "reflected" and "transmitted" (i.e., scattered) by an infinitesimally thin later of water at a particular depth ζ. Note that "reflected" is not synonymous with "backscattered" because downwelling radiance scattered back upward can result from either forward or backward scattering, even for single scattering. Moreover, the "reflectance" described here includes all orders of multiple scattering. Likewise, "transmitted" is not the same as "forward scattered."

The next step of developing the invariant imbedding solution of the VRTE is to formulate *global interaction equations*, which show how light is reflected and transmitted by finitely thick layers of water. In particular, the global interaction equations show how the unknown response radiances (radiances leaving the slab) are related to the known incident radiances falling onto the slab from above and below.

Recall the coordinate system of Fig. 4.1. The water column from just beneath the sea surface at depth $\zeta = w = 0$ to the bottom at depth $\zeta = b$ is conceptually divided into an upper layer from the surface at $\zeta = w$ to an arbitrary depth ζ, and a lower layer from ζ to the bottom at $\zeta = b$. These layers generally have depth dependent IOPs. The layer from the surface to depth ζ will be denoted as the slab $[w, \zeta]$, and the layer from ζ to b is slab $[\zeta, b]$. The entire water column is the union of these two slabs: $[w, b] = [w, \zeta] \cup [\zeta, b]$.

The *linear interaction principle of electromagnetic theory* (Preisendorfer, 1976) *states that the response radiances leaving a layer of water are linear functions of the incident radiances impinging onto the layer from above and below.* As shown in *Light and Water* §8.6, it is possible to develop the global interaction equations beginning with the local interaction equations and developing *fundamental* and *transport* solutions for the radiance amplitudes. However, it is also possible to write down the interaction equations by inspection, given the guidance of Eqs. (4.25) to (4.28) and the corresponding form of the scalar theory seen in *Light and Water*.

[2]The minus signs in the $\underline{\hat{\rho}}_2 \underline{\hat{S}}_2^{\pm}$ and $\underline{\hat{\tau}}_2 \underline{\hat{S}}_2^{\pm}$ terms in Eqs. (4.25) and (4.25) at first seemed some way philosophically wrong to me because they mess up the symmetry between the $\frac{d}{d\zeta}\underline{\hat{S}}_1^{\pm}$ and the $\frac{d}{d\zeta}\underline{\hat{S}}_2^{\pm}$ equations, and because it looks like there are "negative" contributions by reflectance and transmission of the sine amplitudes. However, up to this point the derivation is just straightforward algebra, and the signs are what they are. In any case, minus signs are physically OK because these quantities are all Fourier amplitudes and there is no requirement that a Fourier amplitude "reflectance" for a Fourier amplitude Stokes vector be non-negative. Had I stayed in physical space, there would be no Fourier decomposition, hence no minus signs, and reflectances and transmittances would be more easily interpreted non-negative physical quantities (although the matrices would be much larger).

Consider first the upper layer or slab of water $[w, \zeta]$. Let $\underline{R}(w, \zeta | r, s, u, v)$ denote the matrix that "reflects" or converts downwelling Stokes vectors incident onto the upper surface of the slab at depth w, $\underline{S}^+(w, r, s)$, into upwelling vectors $\underline{S}^-(w, u, v)$ that are leaving the upper surface of the slab. In physical space, the radiance (first element of the Stokes vector) is non-negative, and the corresponding reflectance is a non-negative quantity.

Now, however, the formulation is in terms of Fourier amplitudes, which can be positive or negative even when the physical quantity is positive. Moreover, there are different reflectance operators for cosine and sine amplitudes, as was seen in the local interaction equations. Let $\underline{\underline{\hat{R}}}_1(w, \zeta | \ell)$ denote the matrix that converts downwelling radiance amplitudes incident onto the upper side of the slab, $\underline{\hat{S}}^+_{1,2}(w|\ell)$, into the upwelling radiance amplitudes $\underline{\hat{S}}^-_{1,2}(w|\ell)$ that are leaving the upper surface of the slab, and which leaves radiance cosine (sine) amplitudes as cosine (sine) amplitudes. Let $\underline{\underline{\hat{T}}}_1(w, \zeta | \ell)$ denote the matrix that transmits or converts downwelling radiance amplitudes incident onto the upper surface of the slab at depth w into downwelling radiance amplitudes $\underline{\hat{S}}^+_{1,2}(\zeta|\ell)$ that are leaving the lower surface of the slab at depth ζ. Likewise, let $\underline{\underline{\hat{R}}}_2(w, \zeta | \ell)$ and $\underline{\underline{\hat{T}}}_2(w, \zeta | \ell)$ denote matrices that reflect and transmit radiance amplitudes, but which convert cosine (sine) radiance amplitudes to sine (cosine) amplitudes. Thus $\underline{\underline{\hat{R}}}_1(w, \zeta | \ell)$ functions like the local reflection matrix $\underline{\hat{\rho}}_1(\zeta|\ell)$, $\underline{\underline{\hat{T}}}_2(w, \zeta | \ell)$ functions like $\underline{\hat{\tau}}_2(\zeta|\ell)$, and so on, except that $\underline{\underline{\hat{R}}}_1(w, \zeta | \ell)$ etc. describe the reflectance and transmittance properties of the entire slab of water $[w, \zeta]$. Following this notation, $\underline{\underline{\hat{R}}}_1(\zeta, w | \ell)$ is the matrix the reflects upwelling radiance amplitudes incident onto the slab from below at depth ζ, $\underline{\hat{S}}^-_{1,2}(\zeta|\ell)$, into downwelling radiance amplitudes exiting the slab at ζ; $\underline{\underline{\hat{T}}}_2(\zeta, w | \ell)$ transmits upwelling radiance amplitudes from depth ζ to w and interchanges cosine and sine amplitudes, etc. These $\underline{\underline{\hat{R}}}_1, \underline{\underline{\hat{R}}}_2, \underline{\underline{\hat{T}}}_1$ and $\underline{\underline{\hat{T}}}_2$ matrices are called the (Fourier amplitude versions of the) *standard reflectance* and *standard transmittance* matrices (or operators, or functions) for the slab.

Finally, let $\underline{\hat{\Sigma}}^{-t}_{1,2}(\zeta, w | \ell)$ be the contribution of internal sources to the upwelling radiance amplitudes $\underline{\hat{S}}^-_{1,2}(w|\ell)$ exiting the top of the slab. This radiance is built up within the water column starting at depth ζ and increasing upward. The t in the superscript indicates the *transport* solution, following the notation of *Light and Water*. Similarly, $\underline{\hat{\Sigma}}^{+t}_{1,2}(w, \zeta | \ell)$ is the contribution of internal sources to the downwelling radiance amplitudes $\underline{\hat{S}}^+_{1,2}(\zeta|\ell)$ exiting the bottom of the slab. That radiance is built up within the water column starting at the surface w. Figure 4.3 visually summarizes all of the reflectance and transmittance operations that convert incident radiances into response radiances for the slab $[w, \zeta]$. The notation has been simplified for clarity of the figure: $R_1(w, \zeta)$ represents the matrix $\underline{\underline{\hat{R}}}_1(w, \zeta | \ell)$, etc.

Figure 4.3 shows the various reflectance and transmittance operations. The mathematical statement of those operations must be written to conform to the structure of the local interaction Eqs. (4.25) to (4.28). In particular, terms involving $\underline{\underline{\hat{R}}}_2 \underline{\hat{S}}^{\pm}_2$ and $\underline{\underline{\hat{T}}}_2 \underline{\hat{S}}^{\pm}_2$ must be given minus signs, as seen in Eqs. (4.25) and (4.26)[3]. With this guidance, the global interaction equations for the cosine

[3]I first wrote down (4.29) and (4.30) with plus signs. However, that led to inconsistencies in the ODEs derived below for the standard operators. I then decided–after much frustration–that I have to formulate the global interaction equations with minus signs on the $\underline{\underline{\hat{R}}}_2 \underline{\hat{S}}^{\pm}_2$ and $\underline{\underline{\hat{T}}}_2 \underline{\hat{S}}^{\pm}_2$ terms, just as seen in the $\underline{\hat{\rho}}_2 \underline{\hat{S}}^{\pm}_2$ and $\underline{\hat{\tau}}_2 \underline{\hat{S}}^{\pm}_2$ terms in Eqs. (4.25) and (4.26). I then get consistent ODEs for the standard operators. I have not yet worked out the derivation of the

radiance amplitudes for slab $[w, \zeta]$ (L&W 8.91, 8.92) are

$$\begin{aligned}\hat{\underline{\underline{S}}}_1^-(w|\ell) &= \hat{\underline{\underline{R}}}_1(w,\zeta|\ell)\hat{\underline{\underline{S}}}_1^+(w|\ell) + \hat{\underline{\underline{T}}}_1(\zeta,w|\ell)\hat{\underline{\underline{S}}}_1^-(\zeta|\ell) \\ &\quad - \hat{\underline{\underline{R}}}_2(w,\zeta|\ell)\hat{\underline{\underline{S}}}_2^+(w|\ell) - \hat{\underline{\underline{T}}}_2(\zeta,w|\ell)\hat{\underline{\underline{S}}}_2^-(\zeta|\ell) + \hat{\underline{\underline{\Sigma}}}_1^{-t}(\zeta,w|\ell) \qquad (4.29)\\ \hat{\underline{\underline{S}}}_1^+(\zeta|\ell) &= \hat{\underline{\underline{R}}}_1(\zeta,w|\ell)\hat{\underline{\underline{S}}}_1^-(\zeta|\ell) + \hat{\underline{\underline{T}}}_1(w,\zeta|\ell)\hat{\underline{\underline{S}}}_1^+(w|\ell) \\ &\quad - \hat{\underline{\underline{R}}}_2(\zeta,w|\ell)\hat{\underline{\underline{S}}}_2^-(\zeta|\ell) - \hat{\underline{\underline{T}}}_2(w,\zeta|\ell)\hat{\underline{\underline{S}}}_2^+(w|\ell) + \hat{\underline{\underline{\Sigma}}}_1^{+t}(w,\zeta|\ell) \qquad (4.30)\end{aligned}$$

A similar pair of equations can be written for the sine amplitudes:

$$\begin{aligned}\hat{\underline{\underline{S}}}_2^-(w|\ell) &= \hat{\underline{\underline{R}}}_1(w,\zeta|\ell)\hat{\underline{\underline{S}}}_2^+(w|\ell) + \hat{\underline{\underline{T}}}_1(\zeta,w|\ell)\hat{\underline{\underline{S}}}_2^-(\zeta|\ell) \\ &\quad + \hat{\underline{\underline{R}}}_2(w,\zeta|\ell)\hat{\underline{\underline{S}}}_1^+(w|\ell) + \hat{\underline{\underline{T}}}_2(\zeta,w|\ell)\hat{\underline{\underline{S}}}_1^-(\zeta|\ell) + \hat{\underline{\underline{\Sigma}}}_2^{-t}(\zeta,w|\ell) \qquad (4.31)\\ \hat{\underline{\underline{S}}}_2^+(\zeta|\ell) &= \hat{\underline{\underline{R}}}_1(\zeta,w|\ell)\hat{\underline{\underline{S}}}_2^-(\zeta|\ell) + \hat{\underline{\underline{T}}}_1(w,\zeta|\ell)\hat{\underline{\underline{S}}}_2^+(w|\ell) \\ &\quad + \hat{\underline{\underline{R}}}_2(\zeta,w|\ell)\hat{\underline{\underline{S}}}_1^-(\zeta|\ell) + \hat{\underline{\underline{T}}}_2(w,\zeta|\ell)\hat{\underline{\underline{S}}}_1^+(w|\ell) + \hat{\underline{\underline{\Sigma}}}_2^{+t}(w,\zeta|\ell) \qquad (4.32)\end{aligned}$$

Note that these equations give the response radiances leaving the top and bottom of the slab as functions of the incident radiances. Just as with $\hat{\underline{\underline{\rho}}}_{1,2}$ and $\hat{\underline{\underline{\tau}}}_{1,2}$ in the local interaction equations, the global interaction cosine and sine equations share a common set of standard reflectance and transmittance matrices, but the same matrix operates on difference radiance amplitudes in the cosine and sine pairs of equations.

A tally of the "information content" of the local and global interaction equations is worthwhile. The local interaction Eqs. (4.25)-(4.28) involve 4 matrices of size $4M \times 4M$: $\hat{\underline{\underline{\rho}}}_{1,2}(\zeta|\ell)$ and $\hat{\underline{\underline{\tau}}}_{1,2}(\zeta|\ell)$. However, the corresponding global interaction Eqs. (4.29)-(4.32) involve 8 matrices of size $4M \times 4M$: $\hat{\underline{\underline{R}}}_1(w,\zeta|\ell)$,..., $\hat{\underline{\underline{T}}}_2(\zeta,w|\ell)$. Thus it seems that the global equations some way contain twice as much information as the local equations. This is indeed correct. The local equations contain information about only one depth, ζ. The global equations on the other hand contain information about two depths, w and ζ. Another way to view this is that an infinitesimally thin layer of water reflects and transmits radiance the same for radiance incident from either above or below. However, the IOPs generally depend on depth, in which case finitely thick layers of water do not reflect downwelling radiance from above the layer the same as upwelling radiance from below the layer (ditto for transmission of downwelling vs upwelling radiance). Thus $\hat{\underline{\underline{R}}}_1(w,\zeta|\ell) \neq \hat{\underline{\underline{R}}}_1(\zeta,w|\ell)$, and so on. For finitely thick layers, reflectance and transmittance depend on the direction of the incident light, and there are thus twice as many reflectance and transmittance functions as for an infinitesimal layer.

In one viewpoint, Eqs. (4.29) to 4.32) can be regarded as *definitions* of the standard operators, although they can be derived from the local interaction equations as is done in *Light and Water* §8.6. In any case, the linear interaction principle guarantees the existence of these quantities. Unfortunately, the interaction principle does not specify how to compute these quantities.

global interaction equations starting with the local interaction equations and working through the fundamental and transport solutions, as seen in *Light and Water*, but I assume that this can be done and that the minus signs in the local interaction equations would carry through to the global equations and give the form seen here.

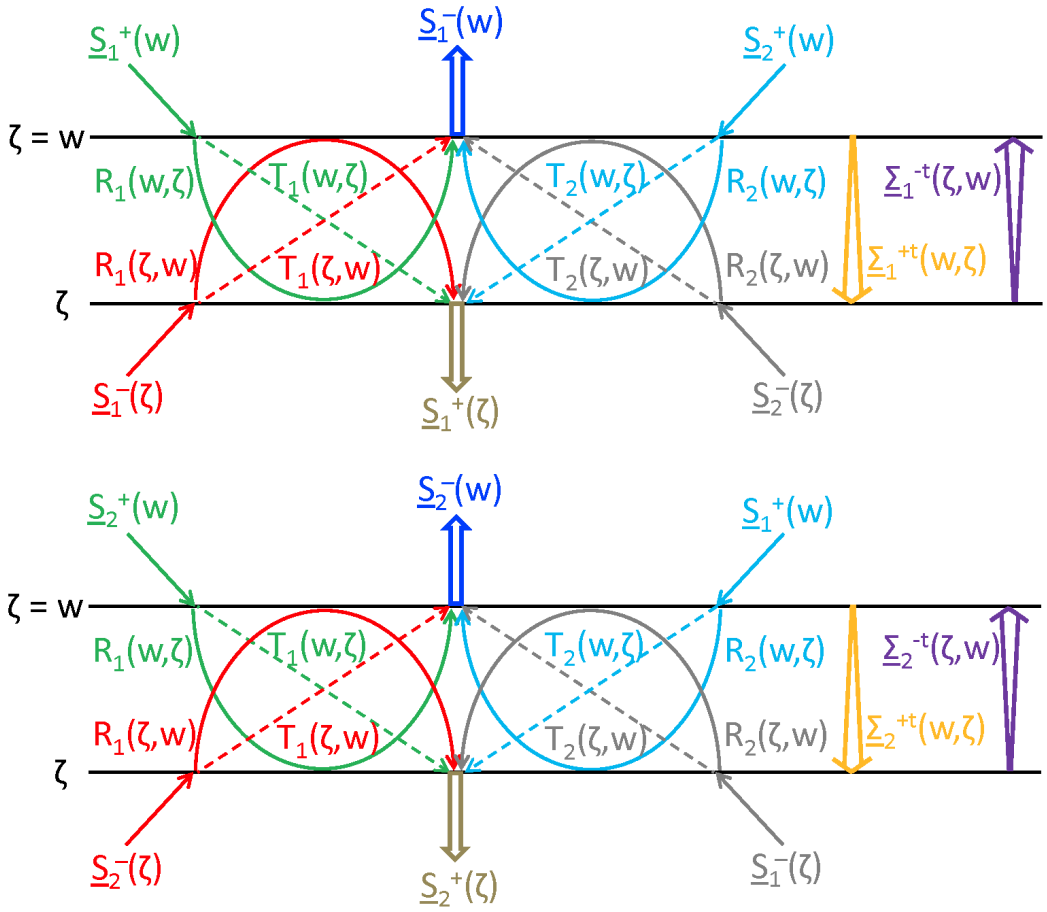

Figure 4.3: Global interaction principles for the slab $[w, \zeta]$. The upper figure is for the cosine amplitudes of the Stokes vector; the lower figure is for the sine amplitudes. Single-width arrows are incident radiances; double arrows are response radiances. Dotted lines represent transmittances, solid lines reflectances, and variable-width arrows internal sources. Colors code the different radiances and the associated reflectance and transmittance operators.

The lower slab $[\zeta, m]$ has a corresponding set of layer standard reflectances and transmittances. These matrices convert radiances incident onto the slab from above at depth ζ and below at depth m into radiances exiting the slab at ζ and m. These operations are illustrated in Fig. 4.4. Note that $\hat{\underline{\underline{S}}}_1^+(\zeta|\ell)$ was a response radiance for slab $[w, \zeta]$, but it is an incident radiance for slab $[\zeta, m]$,

etc. The corresponding global interaction equations are

$$\hat{\underline{S}}_1^-(\zeta|\ell) = \hat{\underline{R}}_1(\zeta,m|\ell)\hat{\underline{S}}_1^+(\zeta|\ell) + \hat{\underline{T}}_1(m,\zeta|\ell)\hat{\underline{S}}_1^-(m|\ell) - \hat{\underline{R}}_2(\zeta,m|\ell)\hat{\underline{S}}_2^+(\zeta|\ell) - \hat{\underline{T}}_2(m,\zeta|\ell)\hat{\underline{S}}_2^-(m|\ell) + \hat{\underline{\Sigma}}_1^{-t}(m,\zeta|\ell) \tag{4.33}$$

$$\hat{\underline{S}}_1^+(m|\ell) = \hat{\underline{R}}_1(m,\zeta|\ell)\hat{\underline{S}}_1^-(m|\ell) + \hat{\underline{T}}_1(\zeta,m|\ell)\hat{\underline{S}}_1^+(\zeta|\ell) - \hat{\underline{R}}_2(m,\zeta|\ell)\hat{\underline{S}}_2^-(m|\ell) - \hat{\underline{T}}_2(\zeta,m|\ell)\hat{\underline{S}}_2^+(\zeta|\ell) + \hat{\underline{\Sigma}}_1^{+t}(\zeta,m|\ell) \tag{4.34}$$

$$\hat{\underline{S}}_2^-(\zeta|\ell) = \hat{\underline{R}}_1(\zeta,m|\ell)\hat{\underline{S}}_2^+(\zeta|\ell) + \hat{\underline{T}}_1(m,\zeta|\ell)\hat{\underline{S}}_2^-(m|\ell) + \hat{\underline{R}}_2(\zeta,m|\ell)\hat{\underline{S}}_1^+(\zeta|\ell) + \hat{\underline{T}}_2(m,\zeta|\ell)\hat{\underline{S}}_1^-(m|\ell) + \hat{\underline{\Sigma}}_2^{-t}(m,\zeta|\ell) \tag{4.35}$$

$$\hat{\underline{S}}_2^+(m|\ell) = \hat{\underline{R}}_1(m,\zeta|\ell)\hat{\underline{S}}_2^-(m|\ell) + \hat{\underline{T}}_1(\zeta,m|\ell)\hat{\underline{S}}_2^+(\zeta|\ell) + \hat{\underline{R}}_2(m,\zeta|\ell)\hat{\underline{S}}_1^-(m|\ell) + \hat{\underline{T}}_2(\zeta,m|\ell)\hat{\underline{S}}_1^+(\zeta|\ell) + \hat{\underline{\Sigma}}_2^{+t}(\zeta,m|\ell) \tag{4.36}$$

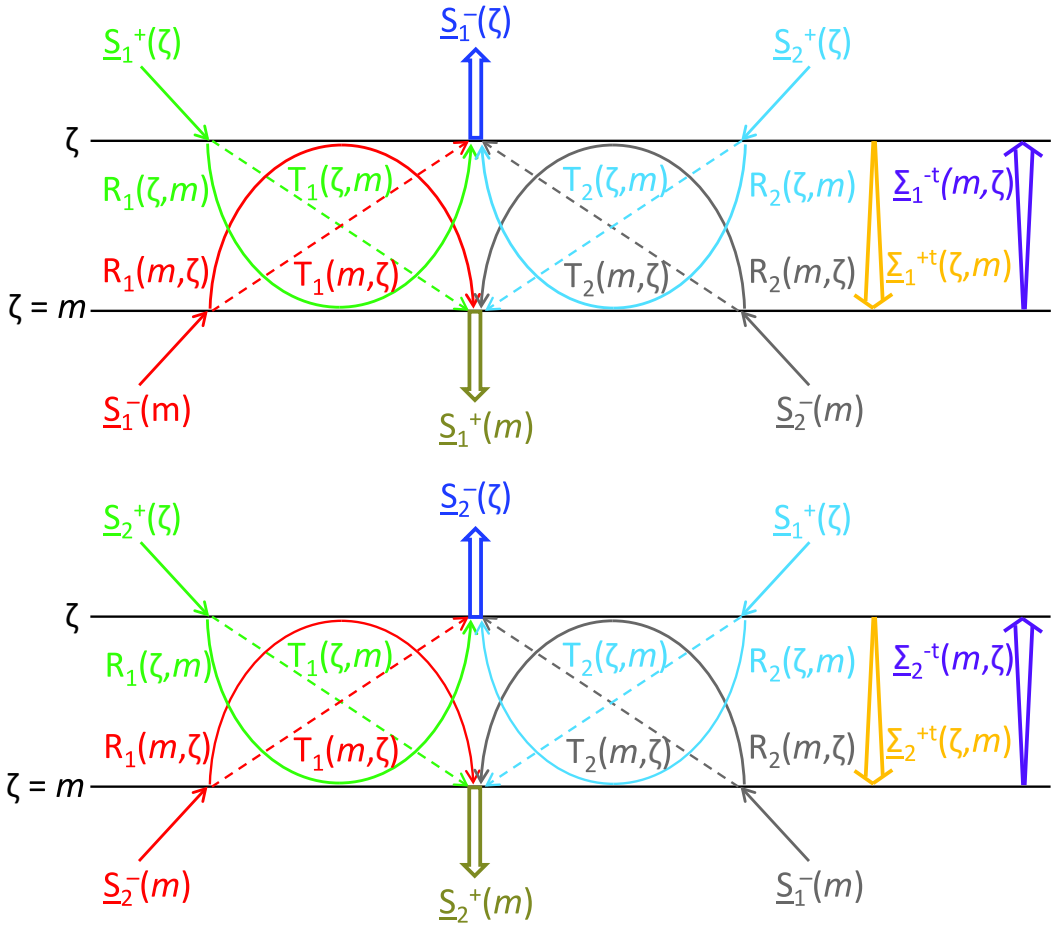

Figure 4.4: Global interaction principles for the slab $[\zeta, m]$.

4.5 Differential Equations for the Standard Matrices

The standard matrices describe how radiance is reflected and transmitted by a layer of water, which in turn is determined by the water absorption and scattering properties. *The standard matrices are therefore spatially integrated inherent optical properties.* The interaction principle guarantees the existence of the standard matrices. Unfortunately, it does not specify how to compute these quantities. The next step in the solution algorithm development is therefore to develop equations that can be solved to obtain the standard matrices, given the water-column IOPs. This is done as follows:

1. The global interaction equations are differentiated with respect to depth.
2. The resulting depth derivatives of the radiance are replaced using the local interaction equations.
3. The global interaction equations are used to replace response radiances by incident radiances.
4. The terms of the resulting equations are grouped into factors multiplying the incident radiances, and a group of terms involving the internal sources.
5. The arbitrariness of the incident radiances and internal sources is used to obtain sets of differential equations for the standard matrices.

4.5.1 Downward Sweep Equations

To illustrate this process, consider, for example, $\underline{\underline{\hat{R}}}_1(\zeta, w|\ell)$. In this development we omit the Fourier mode index ℓ for brevity. This matrix occurs in Eq. (4.30). Differentiating that equation with respect to depth gives

$$\begin{aligned}\frac{d}{d\zeta}\underline{\hat{S}}_1^+(\zeta) = &\left[\frac{d}{d\zeta}\underline{\underline{\hat{R}}}_1(\zeta,w)\right]\underline{\hat{S}}_1^-(\zeta) + \underline{\underline{\hat{R}}}_1(\zeta,w)\frac{d}{d\zeta}\underline{\hat{S}}_1^-(\zeta) + \left[\frac{d}{d\zeta}\underline{\underline{\hat{T}}}_1(w,\zeta)\right]\underline{\hat{S}}_1^+(w)\\ &- \left[\frac{d}{d\zeta}\underline{\underline{\hat{R}}}_2(\zeta,w)\right]\underline{\hat{S}}_2^-(\zeta) - \underline{\underline{\hat{R}}}_2(\zeta,w)\frac{d}{d\zeta}\underline{\hat{S}}_2^-(\zeta) - \left[\frac{d}{d\zeta}\underline{\underline{\hat{T}}}_2(w,\zeta)\right]\underline{\hat{S}}_2^+(w) + \frac{d}{d\zeta}\underline{\hat{\Sigma}}_1^{+t}(w,\zeta)\,.\end{aligned} \tag{4.37}$$

Note that $\underline{\hat{S}}_{1,2}^+(w)$ is the incident radiance at a fixed location w; it therefore does not have a depth derivative. Quantities such as $\underline{\hat{S}}_{1,2}^{\pm}(\zeta)$, on the other hand, have depth derivatives because the lower boundary depth of slab $[w,\zeta]$ is allowed to vary.

The local interaction equations (4.25), (4.26), and (4.28) are next used to replace the three

depth derivatives of the radiances in the preceding equation. The result is

$$\begin{aligned}
&\underline{\underline{\hat{\rho}}}_1\underline{\hat{S}}_1^-(\zeta)+\underline{\underline{\hat{\tau}}}_1\underline{\hat{S}}_1^+(\zeta)-\underline{\underline{\hat{\rho}}}_2\underline{\hat{S}}_2^-(\zeta)-\underline{\underline{\hat{\tau}}}_2\underline{\hat{S}}_2^+(\zeta)+\underline{\hat{\Sigma}}_1^+(\zeta)\\
&\quad=\left[\frac{d}{d\zeta}\underline{\underline{\hat{R}}}_1(\zeta,w)\right]\underline{\hat{S}}_1^-(\zeta)\\
&\quad+\underline{\underline{\hat{R}}}_1(\zeta,w)\left\{-\underline{\underline{\hat{\rho}}}_1\underline{\hat{S}}_1^+(\zeta)-\underline{\underline{\hat{\tau}}}_1\underline{\hat{S}}_1^-(\zeta)+\underline{\underline{\hat{\rho}}}_2\underline{\hat{S}}_2^+(\zeta)+\underline{\underline{\hat{\tau}}}_2\underline{\hat{S}}_2^-(\zeta)-\underline{\hat{\Sigma}}_1^-(\zeta)\right\}\\
&\quad+\left[\frac{d}{d\zeta}\underline{\underline{\hat{T}}}_1(w,\zeta)\right]\underline{\hat{S}}_1^+(w)-\left[\frac{d}{d\zeta}\underline{\underline{\hat{R}}}_2(\zeta,w)\right]\underline{\hat{S}}_2^-(\zeta)\\
&\quad-\underline{\underline{\hat{R}}}_2(\zeta,w)\left\{-\underline{\underline{\hat{\rho}}}_1\underline{\hat{S}}_2^+(\zeta)-\underline{\underline{\hat{\tau}}}_1\underline{\hat{S}}_2^-(\zeta)-\underline{\underline{\hat{\rho}}}_2\underline{\hat{S}}_1^+(\zeta)-\underline{\underline{\hat{\tau}}}_2\underline{\hat{S}}_1^-(\zeta)-\underline{\hat{\Sigma}}_2^-(\zeta)\right\}\\
&\quad-\left[\frac{d}{d\zeta}\underline{\underline{\hat{T}}}_2(w,\zeta)\right]\underline{\hat{S}}_2^+(w)+\frac{d}{d\zeta}\underline{\hat{\Sigma}}_1^{+t}(w,\zeta)
\end{aligned}\tag{4.38}$$

The common $(\zeta|\ell)$ arguments of the $\underline{\underline{\hat{\rho}}}_{1,2}$ and $\underline{\underline{\hat{\tau}}}_{1,2}$ matrices seen in Eqs. (4.25)-(4.28) have been omitted for brevity. However, the depth argument must be retained on the radiances to identify which radiances are incident and which are responses.

The response radiances $\underline{\hat{S}}_{1,2}^+(\zeta)$ in this last equation are now replaced with incident radiances using the global interaction Eqs. (4.30) and (4.32). The terms are then grouped into factors multiplying the incident radiances $\underline{\hat{S}}_{1,2}^+(w)$ and $\underline{\hat{S}}_{1,2}^-(\zeta)$, plus the remaining terms involving the

internal sources. The resulting equation is

$$
\begin{aligned}
0 = & \left\{ \frac{d}{d\zeta}\underline{\underline{\hat{R}}}_1(\zeta,w) - \underline{\hat{\rho}}_1 - \underline{\underline{\hat{R}}}_1(\zeta,w)\underline{\hat{\tau}}_1 + \underline{\underline{\hat{R}}}_2(\zeta,w)\underline{\hat{\tau}}_2 \right. \\
& \qquad - \left[\underline{\hat{\tau}}_1 + \underline{\underline{\hat{R}}}_1(\zeta,w)\underline{\hat{\rho}}_1 - \underline{\underline{\hat{R}}}_2(\zeta,w)\underline{\hat{\rho}}_2\right]\underline{\underline{\hat{R}}}_1(\zeta,w) \\
& \qquad \left. + \left[\underline{\hat{\tau}}_2 + \underline{\underline{\hat{R}}}_1(\zeta,w)\underline{\hat{\rho}}_2 + \underline{\underline{\hat{R}}}_2(\zeta,w)\underline{\hat{\rho}}_1\right]\underline{\underline{\hat{R}}}_2(\zeta,w)\right\}\underline{\hat{S}}_1^-(\zeta) \\
+ & \left\{ - \frac{d}{d\zeta}\underline{\underline{\hat{R}}}_2(\zeta,w) + \underline{\hat{\rho}}_2 + \underline{\underline{\hat{R}}}_1(\zeta,w)\underline{\hat{\tau}}_2 + \underline{\underline{\hat{R}}}_2(\zeta,w)\underline{\hat{\tau}}_1 \right. \\
& \qquad + \left[\underline{\hat{\tau}}_1 + \underline{\underline{\hat{R}}}_1(\zeta,w)\underline{\hat{\rho}}_1 - \underline{\underline{\hat{R}}}_2(\zeta,w)\underline{\hat{\rho}}_2\right]\underline{\underline{\hat{R}}}_2(\zeta,w) \\
& \qquad \left. + \left[\underline{\hat{\tau}}_2 + \underline{\underline{\hat{R}}}_1(\zeta,w)\underline{\hat{\rho}}_2 + \underline{\underline{\hat{R}}}_2(\zeta,w)\underline{\hat{\rho}}_1\right]\underline{\underline{\hat{R}}}_1(\zeta,w)\right\}\underline{\hat{S}}_2^-(\zeta) \\
& \left\{ \frac{d}{d\zeta}\underline{\underline{\hat{T}}}_1(w,\zeta) - \left[\underline{\hat{\tau}}_1 + \underline{\underline{\hat{R}}}_1(\zeta,w)\underline{\hat{\rho}}_1 - \underline{\underline{\hat{R}}}_2(\zeta,w)\underline{\hat{\rho}}_2\right]\underline{\underline{\hat{T}}}_1(w,\zeta) \right. \\
& \qquad \left. + \left[\underline{\hat{\tau}}_2 + \underline{\underline{\hat{R}}}_1(\zeta,w)\underline{\hat{\rho}}_2 + \underline{\underline{\hat{R}}}_2(\zeta,w)\underline{\hat{\rho}}_1\right]\underline{\underline{\hat{T}}}_2(w,\zeta)\right\}\underline{\hat{S}}_1^+(w) \\
+ & \left\{ - \frac{d}{d\zeta}\underline{\underline{\hat{T}}}_2(w,\zeta) + \left[\underline{\hat{\tau}}_1 + \underline{\underline{\hat{R}}}_1(\zeta,w)\underline{\hat{\rho}}_1 - \underline{\underline{\hat{R}}}_2(\zeta,w)\underline{\hat{\rho}}_2\right]\underline{\underline{\hat{T}}}_2(w,\zeta) \right. \\
& \qquad \left. + \left[\underline{\hat{\tau}}_2 + \underline{\underline{\hat{R}}}_1(\zeta,w)\underline{\hat{\rho}}_2 + \underline{\underline{\hat{R}}}_2(\zeta,w)\underline{\hat{\rho}}_1\right]\underline{\underline{\hat{T}}}_1(w,\zeta)\right\}\underline{\hat{S}}_2^+(w) \\
& + \left\{ \frac{d}{d\zeta}\underline{\hat{\Sigma}}_1^{+t}(w,\zeta) - \underline{\hat{\Sigma}}_1^+(\zeta) - \underline{\underline{\hat{R}}}_1(\zeta,w)\underline{\hat{\Sigma}}_1^-(\zeta) + \underline{\underline{\hat{R}}}_2(\zeta,w)\underline{\hat{\Sigma}}_2^-(\zeta) \right. \\
& \qquad - \left[\underline{\hat{\tau}}_1 + \underline{\underline{\hat{R}}}_1(\zeta,w)\underline{\hat{\rho}}_1 - \underline{\underline{\hat{R}}}_2(\zeta,w)\underline{\hat{\rho}}_2\right]\underline{\hat{\Sigma}}_1^{+t}(w,\zeta) \\
& \qquad \left. + \left[\underline{\hat{\tau}}_2 + \underline{\underline{\hat{R}}}_1(\zeta,w)\underline{\hat{\rho}}_2 + \underline{\underline{\hat{R}}}_2(\zeta,w)\underline{\hat{\rho}}_1\right]\underline{\hat{\Sigma}}_2^{+t}(w,\zeta)\right\}
\end{aligned}
\tag{4.39}
$$

The first four groups of terms in brackets multiply incident radiances. The last group of terms contains the internal sources.

The four incident radiances and the internal sources have arbitrary values that must be specified as known inputs to the radiative transfer problem. For example, all of the input radiances except for $\underline{\hat{S}}_1^+(w)$ can be zero, and all of the internal sources can be zero. In that case, the terms in brackets multiplying $\underline{\hat{S}}_1^+(w)$ must sum to zero. Similarly, all of the incident radiances can be zero, but there can be internal sources generating the light field within the water. In that case, the last group of bracketed terms must sum to zero. Arguments of this type show that each group of terms in brackets must be zero. Setting the individual groups of bracketed terms to zero then gives a set

of five coupled nonlinear ordinary differential equations (ODEs) (L&W 8.74, 8.75, and 8.78):

$$\begin{aligned}
\frac{d}{d\zeta}\underline{\underline{\hat{R}}}_1(\zeta,w) &= \underline{\hat{\rho}}_1 + \underline{\underline{\hat{R}}}_1(\zeta,w)\underline{\hat{\tau}}_1 - \underline{\underline{\hat{R}}}_2(\zeta,w)\underline{\hat{\tau}}_2 \\
&+ \left[\underline{\hat{\tau}}_1 + \underline{\underline{\hat{R}}}_1(\zeta,w)\underline{\hat{\rho}}_1 - \underline{\underline{\hat{R}}}_2(\zeta,w)\underline{\hat{\rho}}_2\right]\underline{\underline{\hat{R}}}_1(\zeta,w) \\
&- \left[\underline{\hat{\tau}}_2 + \underline{\underline{\hat{R}}}_1(\zeta,w)\underline{\hat{\rho}}_2 + \underline{\underline{\hat{R}}}_2(\zeta,w)\underline{\hat{\rho}}_1\right]\underline{\underline{\hat{R}}}_2(\zeta,w) \qquad (4.40) \\
\frac{d}{d\zeta}\underline{\underline{\hat{R}}}_2(\zeta,w) &= \underline{\hat{\rho}}_2 + \underline{\underline{\hat{R}}}_1(\zeta,w)\underline{\hat{\tau}}_2 + \underline{\underline{\hat{R}}}_2(\zeta,w)\underline{\hat{\tau}}_1 \\
&+ \left[\underline{\hat{\tau}}_1 + \underline{\underline{\hat{R}}}_1(\zeta,w)\underline{\hat{\rho}}_1 - \underline{\underline{\hat{R}}}_2(\zeta,w)\underline{\hat{\rho}}_2\right]\underline{\underline{\hat{R}}}_2(\zeta,w) \\
&+ \left[\underline{\hat{\tau}}_2 + \underline{\underline{\hat{R}}}_1(\zeta,w)\underline{\hat{\rho}}_2 + \underline{\underline{\hat{R}}}_2(\zeta,w)\underline{\hat{\rho}}_1\right]\underline{\underline{\hat{R}}}_1(\zeta,w) \qquad (4.41) \\
\frac{d}{d\zeta}\underline{\underline{\hat{T}}}_1(w,\zeta) &= \left[\underline{\hat{\tau}}_1 + \underline{\underline{\hat{R}}}_1(\zeta,w)\underline{\hat{\rho}}_1 - \underline{\underline{\hat{R}}}_2(\zeta,w)\underline{\hat{\rho}}_2\right]\underline{\underline{\hat{T}}}_1(w,\zeta) \\
&- \left[\underline{\hat{\tau}}_2 + \underline{\underline{\hat{R}}}_1(\zeta,w)\underline{\hat{\rho}}_2 + \underline{\underline{\hat{R}}}_2(\zeta,w)\underline{\hat{\rho}}_1\right]\underline{\underline{\hat{T}}}_2(w,\zeta) \qquad (4.42) \\
\frac{d}{d\zeta}\underline{\underline{\hat{T}}}_2(w,\zeta) &= \left[\underline{\hat{\tau}}_1 + \underline{\underline{\hat{R}}}_1(\zeta,w)\underline{\hat{\rho}}_1 - \underline{\underline{\hat{R}}}_2(\zeta,w)\underline{\hat{\rho}}_2\right]\underline{\underline{\hat{T}}}_2(w,\zeta) \\
&+ \left[\underline{\hat{\tau}}_2 + \underline{\underline{\hat{R}}}_1(\zeta,w)\underline{\hat{\rho}}_2 + \underline{\underline{\hat{R}}}_2(\zeta,w)\underline{\hat{\rho}}_1\right]\underline{\underline{\hat{T}}}_1(w,\zeta) \qquad (4.43) \\
\frac{d}{d\zeta}\underline{\underline{\hat{\Sigma}}}_1^{+t}(w,\zeta) &= \underline{\underline{\hat{\Sigma}}}_1^{+}(\zeta) + \underline{\underline{\hat{R}}}_1(\zeta,w)\underline{\underline{\hat{\Sigma}}}_1^{-}(\zeta) - \underline{\underline{\hat{R}}}_2(\zeta,w)\underline{\underline{\hat{\Sigma}}}_2^{-}(\zeta) \\
&+ \left[\underline{\hat{\tau}}_1 + \underline{\underline{\hat{R}}}_1(\zeta,w)\underline{\hat{\rho}}_1 - \underline{\underline{\hat{R}}}_2(\zeta,w)\underline{\hat{\rho}}_2\right]\underline{\underline{\hat{\Sigma}}}_1^{+t}(w,\zeta) \\
&- \left[\underline{\hat{\tau}}_2 + \underline{\underline{\hat{R}}}_1(\zeta,w)\underline{\hat{\rho}}_2 + \underline{\underline{\hat{R}}}_2(\zeta,w)\underline{\hat{\rho}}_1\right]\underline{\underline{\hat{\Sigma}}}_2^{+t}(w,\zeta) \qquad (4.44)
\end{aligned}$$

This self-contained set of ODEs can be solved given the IOPs $\underline{\hat{\rho}}_{1,2}(\zeta)$, $\underline{\hat{\tau}}_{1,2}(\zeta)$, and initial conditions. The initial conditions are deduced as follows. If $\zeta \to w = 0$, there will be no reflectance by a slab $[w, w]$ of zero thickness, and all of the radiance will be transmitted. Likewise, there will be no contribution by internal sources in a slab of zero thickness. Thus the initial conditions at $\zeta = w$ are simply

$$\underline{\underline{\hat{R}}}_1(w,w) = \underline{\underline{\hat{R}}}_2(w,w) = \underline{\underline{0}}_{4M\times 4M} \qquad (4.45)$$

$$\underline{\underline{\hat{T}}}_1(w,w) = \underline{\underline{\hat{T}}}_2(w,w) = \underline{\underline{I}}_{4M\times 4M} \qquad (4.46)$$

$$\underline{\underline{\hat{\Sigma}}}_1^{+t}(w,w) = \underline{\underline{0}}_{4M\times 1} \qquad (4.47)$$

where $\underline{\underline{0}}_{4M\times 4M}$ is a $4M \times 4M$ matrix of zeros, $\underline{\underline{I}}_{4M\times 4M}$ is a $4M \times 4M$ identity matrix, and $\underline{\underline{0}}_{4M\times 1}$ is a $4M \times 1$ column vector of zeros.

Equations (4.40) to (4.44) can therefore be integrated in a "downward sweep" from depth $w = 0$ to any depth ζ, starting with the initial conditions of Eqs. (4.45) to (4.47).

It should be noted that the reflectance and transmittance properties of the slab $[w,\zeta]$ *can be computed without explicit consideration of the air-water surface* $[a,w]$. Slab $[w,\zeta]$ is therefore referred to as a "bare" slab, i.e., a water body without an air-water surface. This is physically equivalent to a water body with the same index of refraction as the air, which would not reflect or refract light passing though such a surface. The effects of the actual air-sea surface will be incorporated into the solution in the next chapter.

Differentiating global interaction Eq. (4.29) and repeating this process gives the following set of equations (L&W 8.76, 8.77, 8.78):

$$\begin{aligned}
\frac{d}{d\zeta}\underline{\underline{\hat{R}}}_1(w,\zeta) &= \left[\underline{\underline{\hat{T}}}_1(\zeta,w)\underline{\underline{\hat{\rho}}}_1 - \underline{\underline{\hat{T}}}_2(\zeta,w)\underline{\underline{\hat{\rho}}}_2\right]\underline{\underline{\hat{T}}}_1(w,\zeta) \\
&- \left[\underline{\underline{\hat{T}}}_1(\zeta,w)\underline{\underline{\hat{\rho}}}_2 + \underline{\underline{\hat{T}}}_2(\zeta,w)\underline{\underline{\hat{\rho}}}_1\right]\underline{\underline{\hat{T}}}_2(w,\zeta) \qquad (4.48)\\
\frac{d}{d\zeta}\underline{\underline{\hat{R}}}_2(w,\zeta) &= \left[\underline{\underline{\hat{T}}}_1(\zeta,w)\underline{\underline{\hat{\rho}}}_1 - \underline{\underline{\hat{T}}}_2(\zeta,w)\underline{\underline{\hat{\rho}}}_2\right]\underline{\underline{\hat{T}}}_2(w,\zeta) \\
&+ \left[\underline{\underline{\hat{T}}}_1(\zeta,w)\underline{\underline{\hat{\rho}}}_2 + \underline{\underline{\hat{T}}}_2(\zeta,w)\underline{\underline{\hat{\rho}}}_1\right]\underline{\underline{\hat{T}}}_1(w,\zeta) \qquad (4.49)\\
\frac{d}{d\zeta}\underline{\underline{\hat{T}}}_1(\zeta,w) &= \underline{\underline{\hat{T}}}_1(\zeta,w)\underline{\hat{\tau}}_1 - \underline{\underline{\hat{T}}}_2(\zeta,w)\underline{\hat{\tau}}_2 \\
&+ \left[\underline{\underline{\hat{T}}}_1(\zeta,w)\underline{\underline{\hat{\rho}}}_1 - \underline{\underline{\hat{T}}}_2(\zeta,w)\underline{\underline{\hat{\rho}}}_2\right]\underline{\underline{\hat{R}}}_1(\zeta,w) \\
&- \left[\underline{\underline{\hat{T}}}_1(\zeta,w)\underline{\underline{\hat{\rho}}}_2 + \underline{\underline{\hat{T}}}_2(\zeta,w)\underline{\underline{\hat{\rho}}}_1\right]\underline{\underline{\hat{R}}}_2(\zeta,w) \qquad (4.50)\\
\frac{d}{d\zeta}\underline{\underline{\hat{T}}}_2(\zeta,w) &= \underline{\underline{\hat{T}}}_1(\zeta,w)\underline{\hat{\tau}}_2 + \underline{\underline{\hat{T}}}_2(\zeta,w)\underline{\hat{\tau}}_1 \\
&+ \left[\underline{\underline{\hat{T}}}_1(\zeta,w)\underline{\underline{\hat{\rho}}}_1 - \underline{\underline{\hat{T}}}_2(\zeta,w)\underline{\underline{\hat{\rho}}}_2\right]\underline{\underline{\hat{R}}}_2(\zeta,w) \\
&+ \left[\underline{\underline{\hat{T}}}_1(\zeta,w)\underline{\underline{\hat{\rho}}}_2 + \underline{\underline{\hat{T}}}_2(\zeta,w)\underline{\underline{\hat{\rho}}}_1\right]\underline{\underline{\hat{R}}}_1(\zeta,w) \qquad (4.51)\\
\frac{d}{d\zeta}\underline{\underline{\hat{\Sigma}}}_1^{-t}(\zeta,w) &= \underline{\underline{\hat{T}}}_1(\zeta,w)\underline{\underline{\hat{\Sigma}}}_1^{-}(\zeta) - \underline{\underline{\hat{T}}}_2(\zeta,w)\underline{\underline{\hat{\Sigma}}}_2^{-}(\zeta)) \\
&+ \left[\underline{\underline{\hat{T}}}_1(\zeta,w)\underline{\underline{\hat{\rho}}}_1 - \underline{\underline{\hat{T}}}_2(\zeta,w)\underline{\underline{\hat{\rho}}}_2\right]\underline{\underline{\hat{\Sigma}}}_1^{+t}(w,\zeta) \\
&- \left[\underline{\underline{\hat{T}}}_1(\zeta,w)\underline{\underline{\hat{\rho}}}_2 + \underline{\underline{\hat{T}}}_2(\zeta,w)\underline{\underline{\hat{\rho}}}_1\right]\underline{\underline{\hat{\Sigma}}}_2^{+t}(w,\zeta) \qquad (4.52)
\end{aligned}$$

The preceding development gives ODEs for all of the reflectance and transmittance matrices seen in Fig. 4.3, and for the cosine amplitudes of the internal source terms. Equations for the remaining internal source sine amplitudes can be obtained from Eqs. (4.31) and (4.32). Differentiating Eq. (4.31) and repeating the above process gives the equation for $\underline{\underline{\hat{\Sigma}}}_2^{-t}(\zeta, w)$; starting with Eq. (4.32) gives the ODE for $\underline{\underline{\hat{\Sigma}}}_2^{+t}(\zeta, w)$. Those derivations also re-derive the equations for $\underline{\underline{R}}_{1,2}(\zeta, w)$, $\underline{\underline{T}}_{1,2}(w, \zeta)$, $\underline{\underline{R}}_{1,2}(w, \zeta)$ and $\underline{\underline{T}}_{1,2}(\zeta, w)$, which serves as a check on those equations[4]. The results for the remaining source terms are

$$\begin{aligned}
\frac{d}{d\zeta}\underline{\underline{\hat{\Sigma}}}_2^{-t}(\zeta, w) &= \underline{\underline{\hat{T}}}_1(\zeta, w)\underline{\underline{\hat{\Sigma}}}_2^{-}(\zeta) + \underline{\underline{\hat{T}}}_2(\zeta, w)\underline{\underline{\hat{\Sigma}}}_1^{-}(\zeta) \\
&+ \left[\underline{\underline{\hat{T}}}_1(\zeta, w)\underline{\hat{\rho}}_2 + \underline{\underline{\hat{T}}}_2(\zeta, w)\underline{\hat{\rho}}_1\right]\underline{\underline{\hat{\Sigma}}}_1^{+t}(w, \zeta) \\
&+ \left[\underline{\underline{\hat{T}}}_1(\zeta, w)\underline{\hat{\rho}}_1 - \underline{\underline{\hat{T}}}_2(\zeta, w)\underline{\hat{\rho}}_2\right]\underline{\underline{\hat{\Sigma}}}_2^{+t}(w, \zeta) \qquad (4.53) \\
\frac{d}{d\zeta}\underline{\underline{\hat{\Sigma}}}_2^{+t}(\zeta, w) &= \underline{\underline{\hat{\Sigma}}}_2^{+}(\zeta) + \underline{\underline{\hat{R}}}_1(\zeta, w)\underline{\underline{\hat{\Sigma}}}_2^{-}(\zeta) + \underline{\underline{\hat{R}}}_2(\zeta, w)\underline{\underline{\hat{\Sigma}}}_1^{-}(\zeta)) \\
&+ \left[\underline{\hat{\tau}}_2 + \underline{\underline{\hat{R}}}_1(\zeta, w)\underline{\hat{\rho}}_2 + \underline{\underline{\hat{R}}}_2(\zeta, w)\underline{\hat{\rho}}_1\right]\underline{\underline{\hat{\Sigma}}}_1^{+t}(w, \zeta) \\
&+ \left[\underline{\hat{\tau}}_1 + \underline{\underline{\hat{R}}}_1(\zeta, w)\underline{\hat{\rho}}_1 - \underline{\underline{\hat{R}}}_2(\zeta, w)\underline{\hat{\rho}}_2\right]\underline{\underline{\hat{\Sigma}}}_1^{+t}(w, \zeta) \qquad (4.54)
\end{aligned}$$

These equations can be integrated from $\zeta = w$ downward, starting with the initial conditions

$$\underline{\underline{\hat{R}}}_1(w, w) = \underline{\underline{\hat{R}}}_2(w, w) = \underline{\underline{0}}_{4M\times 4M} \qquad (4.55)$$

$$\underline{\underline{\hat{T}}}_1(w, w) = \underline{\underline{\hat{T}}}_2(w, w) = \underline{\underline{I}}_{4M\times 4M} \qquad (4.56)$$

$$\underline{\underline{\hat{\Sigma}}}_1^{-t}(w, w) = \underline{\underline{0}}_{4M\times 1} \qquad (4.57)$$

$$\underline{\underline{\hat{\Sigma}}}_2^{\pm t}(w, w) = \underline{\underline{0}}_{4M\times 1} \qquad (4.58)$$

Note that the set of Eqs. (4.40)-(4.44) is self contained, whereas Eqs. (4.48)-(4.52) contain terms from the first set. The set (4.40)-(4.44) is therefore called the "major" set, and (4.48)-(4.52) the "minor" set. Equation (4.54) belongs to the major set, and (4.53) to the minor set. In any case, these equations must be solved simultaneously.

[4]Before I inserted the minus signs into the global interaction Eqs. (4.29) and (4.30), the derivation of Eqs. (4.53) and (4.54) starting with (4.32) and (4.31) led to different equations for $\underline{\underline{R}}_{1,2}(\zeta, w)$, $\underline{\underline{T}}_{1,2}(w, \zeta)$, $\underline{\underline{R}}_{1,2}(w, \zeta)$ and $\underline{\underline{T}}_{1,2}(\zeta, w)$ than what I got starting with (4.29) and (4.30) (the minus signs appeared in different places than before). This is what led me to put the minus signs into the global interaction equations. The two paths to the ODEs for $\underline{\underline{R}}_{1,2}(\zeta, w)$ etc. then give the same equations regardless of which global interaction equation is used as the starting point.

4.5.2 Upward Sweep Equations

The slab $[\zeta, m]$ illustrated in Fig. 4.4 has a corresponding set of ODEs for the standard matrices shown in the figure. The derivations follow the same procedure as for slab $[w, \zeta]$, so the details do not need to be presented. The resulting equations are as follows:
The major set, obtained starting with Eq. (4.33) and (4.35):

$$
\begin{aligned}
-\frac{d}{d\zeta}\hat{\underline{\underline{R}}}_1(\zeta,m) &= \hat{\underline{\rho}}_1 + \hat{\underline{\underline{R}}}_1(\zeta,m)\hat{\underline{\tau}}_1 - \hat{\underline{\underline{R}}}_2(\zeta,m)\hat{\underline{\tau}}_2 \\
&+ \left[\hat{\underline{\tau}}_1 + \hat{\underline{\underline{R}}}_1(\zeta,m)\hat{\underline{\rho}}_1 - \hat{\underline{\underline{R}}}_2(\zeta,m)\hat{\underline{\rho}}_2\right]\hat{\underline{\underline{R}}}_1(\zeta,m) \\
&- \left[\hat{\underline{\tau}}_2 + \hat{\underline{\underline{R}}}_1(\zeta,m)\hat{\underline{\rho}}_2 + \hat{\underline{\underline{R}}}_2(\zeta,m)\hat{\underline{\rho}}_1\right]\hat{\underline{\underline{R}}}_2(\zeta,m) \qquad (4.59)\\
-\frac{d}{d\zeta}\hat{\underline{\underline{R}}}_2(\zeta,m) &= \hat{\underline{\rho}}_2 + \hat{\underline{\underline{R}}}_1(\zeta,m)\hat{\underline{\tau}}_2 + \hat{\underline{\underline{R}}}_2(\zeta,m)\hat{\underline{\tau}}_1 \\
&+ \left[\hat{\underline{\tau}}_1 + \hat{\underline{\underline{R}}}_1(\zeta,m)\hat{\underline{\rho}}_1 - \hat{\underline{\underline{R}}}_2(\zeta,m)\hat{\underline{\rho}}_2\right]\hat{\underline{\underline{R}}}_2(\zeta,m) \\
&+ \left[\hat{\underline{\tau}}_2 + \hat{\underline{\underline{R}}}_1(\zeta,m)\hat{\underline{\rho}}_2 + \hat{\underline{\underline{R}}}_2(\zeta,m)\hat{\underline{\rho}}_1\right]\hat{\underline{\underline{R}}}_1(\zeta,m) \qquad (4.60)\\
-\frac{d}{d\zeta}\hat{\underline{\underline{T}}}_1(m,\zeta) &= \left[\hat{\underline{\tau}}_1 + \hat{\underline{\underline{R}}}_1(\zeta,m)\hat{\underline{\rho}}_1 - \hat{\underline{\underline{R}}}_2(\zeta,m)\hat{\underline{\rho}}_2\right]\hat{\underline{\underline{T}}}_1(m,\zeta) \\
&- \left[\hat{\underline{\tau}}_2 + \hat{\underline{\underline{R}}}_1(\zeta,m)\hat{\underline{\rho}}_2 + \hat{\underline{\underline{R}}}_2(\zeta,m)\hat{\underline{\rho}}_1\right]\hat{\underline{\underline{T}}}_2(m,\zeta) \qquad (4.61)\\
-\frac{d}{d\zeta}\hat{\underline{\underline{T}}}_2(m,\zeta) &= \left[\hat{\underline{\tau}}_1 + \hat{\underline{\underline{R}}}_1(\zeta,m)\hat{\underline{\rho}}_1 - \hat{\underline{\underline{R}}}_2(\zeta,m)\hat{\underline{\rho}}_2\right]\hat{\underline{\underline{T}}}_2(m,\zeta) \\
&+ \left[\hat{\underline{\tau}}_2 + \hat{\underline{\underline{R}}}_1(\zeta,m)\hat{\underline{\rho}}_2 + \hat{\underline{\underline{R}}}_2(\zeta,m)\hat{\underline{\rho}}_1\right]\hat{\underline{\underline{T}}}_1(m,\zeta) \qquad (4.62)\\
-\frac{d}{d\zeta}\hat{\underline{\underline{\Sigma}}}_1^{-t}(m,\zeta) &= \hat{\underline{\underline{\Sigma}}}_1^{-}(\zeta) + \hat{\underline{\underline{R}}}_1(\zeta,m)\hat{\underline{\underline{\Sigma}}}_1^{+}(\zeta) - \hat{\underline{\underline{R}}}_2(\zeta,m)\hat{\underline{\underline{\Sigma}}}_2^{+}(\zeta) \\
&+ \left[\hat{\underline{\tau}}_1 + \hat{\underline{\underline{R}}}_1(\zeta,m)\hat{\underline{\rho}}_1 - \hat{\underline{\underline{R}}}_2(\zeta,m)\hat{\underline{\rho}}_2\right]\hat{\underline{\underline{\Sigma}}}_1^{-t}(m,\zeta) \\
&- \left[\hat{\underline{\tau}}_2 + \hat{\underline{\underline{R}}}_1(\zeta,m)\hat{\underline{\rho}}_2 + \hat{\underline{\underline{R}}}_2(\zeta,m)\hat{\underline{\rho}}_1\right]\hat{\underline{\underline{\Sigma}}}_2^{-t}(m,\zeta) \qquad (4.63)\\
-\frac{d}{d\zeta}\hat{\underline{\underline{\Sigma}}}_2^{-t}(m,\zeta) &= \hat{\underline{\underline{\Sigma}}}_2^{-}(\zeta) + \hat{\underline{\underline{R}}}_1(\zeta,m)\hat{\underline{\underline{\Sigma}}}_2^{+}(\zeta) + \hat{\underline{\underline{R}}}_2(\zeta,m)\hat{\underline{\underline{\Sigma}}}_1^{+}(\zeta)) \\
&+ \left[\hat{\underline{\tau}}_1 + \hat{\underline{\underline{R}}}_1(\zeta,m)\hat{\underline{\rho}}_1 - \hat{\underline{\underline{R}}}_2(\zeta,m)\hat{\underline{\rho}}_2\right]\hat{\underline{\underline{\Sigma}}}_2^{-t}(m,\zeta) \\
&+ \left[\hat{\underline{\tau}}_2 + \hat{\underline{\underline{R}}}_1(\zeta,m)\hat{\underline{\rho}}_2 + \hat{\underline{\underline{R}}}_2(\zeta,m)\hat{\underline{\rho}}_1\right]\hat{\underline{\underline{\Sigma}}}_1^{-t}(m,\zeta) \qquad (4.64)
\end{aligned}
$$

The minor set, obtained starting with Eqs. (4.34) and (4.36) :

$$
\begin{aligned}
-\frac{d}{d\zeta}\hat{\underline{\underline{R}}}_1(m,\zeta) &= \left[\hat{\underline{\underline{T}}}_1(\zeta,m)\hat{\underline{\underline{\rho}}}_1 - \hat{\underline{\underline{T}}}_2(\zeta,m)\hat{\underline{\underline{\rho}}}_2\right]\hat{\underline{\underline{T}}}_1(m,\zeta) \\
&- \left[\hat{\underline{\underline{T}}}_1(\zeta,m)\hat{\underline{\underline{\rho}}}_2 + \hat{\underline{\underline{T}}}_2(\zeta,m)\hat{\underline{\underline{\rho}}}_1\right]\hat{\underline{\underline{T}}}_2(m,\zeta) \qquad (4.65)\\
-\frac{d}{d\zeta}\hat{\underline{\underline{R}}}_2(m,\zeta) &= \left[\hat{\underline{\underline{T}}}_1(\zeta,m)\hat{\underline{\underline{\rho}}}_1 - \hat{\underline{\underline{T}}}_2(\zeta,m)\hat{\underline{\underline{\rho}}}_2\right]\hat{\underline{\underline{T}}}_2(m,\zeta) \\
&+ \left[\hat{\underline{\underline{T}}}_1(\zeta,m)\hat{\underline{\underline{\rho}}}_2 + \hat{\underline{\underline{T}}}_2(\zeta,m)\hat{\underline{\underline{\rho}}}_1\right]\hat{\underline{\underline{T}}}_1(m,\zeta) \qquad (4.66)\\
-\frac{d}{d\zeta}\hat{\underline{\underline{T}}}_1(\zeta,m) &= \hat{\underline{\underline{T}}}_1(\zeta,m)\hat{\underline{\underline{\tau}}}_1 - \hat{\underline{\underline{T}}}_2(\zeta,m)\hat{\underline{\underline{\tau}}}_2 \\
&+ \left[\hat{\underline{\underline{T}}}_1(\zeta,m)\hat{\underline{\underline{\rho}}}_1 - \hat{\underline{\underline{T}}}_2(\zeta,m)\hat{\underline{\underline{\rho}}}_2\right]\hat{\underline{\underline{R}}}_1(\zeta,m) \\
&- \left[\hat{\underline{\underline{T}}}_1(\zeta,m)\hat{\underline{\underline{\rho}}}_2 + \hat{\underline{\underline{T}}}_2(\zeta,m)\hat{\underline{\underline{\rho}}}_1\right]\hat{\underline{\underline{R}}}_2(\zeta,m) \qquad (4.67)\\
-\frac{d}{d\zeta}\hat{\underline{\underline{T}}}_2(\zeta,m) &= \hat{\underline{\underline{T}}}_1(\zeta,m)\hat{\underline{\underline{\tau}}}_2 + \hat{\underline{\underline{T}}}_2(\zeta,m)\hat{\underline{\underline{\tau}}}_1 \\
&+ \left[\hat{\underline{\underline{T}}}_1(\zeta,m)\hat{\underline{\underline{\rho}}}_1 - \hat{\underline{\underline{T}}}_2(\zeta,m)\hat{\underline{\underline{\rho}}}_2\right]\hat{\underline{\underline{R}}}_2(\zeta,m) \\
&+ \left[\hat{\underline{\underline{T}}}_1(\zeta,m)\hat{\underline{\underline{\rho}}}_2 + \hat{\underline{\underline{T}}}_2(\zeta,m)\hat{\underline{\underline{\rho}}}_1\right]\hat{\underline{\underline{R}}}_1(\zeta,m) \qquad (4.68)\\
-\frac{d}{d\zeta}\hat{\underline{\underline{\Sigma}}}_1^{+t}(\zeta,m) &= \hat{\underline{\underline{T}}}_1(\zeta,m)\hat{\underline{\underline{\Sigma}}}_1^{+}(\zeta) - \hat{\underline{\underline{T}}}_2(\zeta,m)\hat{\underline{\underline{\Sigma}}}_2^{+}(\zeta)) \\
&+ \left[\hat{\underline{\underline{T}}}_1(\zeta,m)\hat{\underline{\underline{\rho}}}_1 - \hat{\underline{\underline{T}}}_2(\zeta,m)\hat{\underline{\underline{\rho}}}_2\right]\hat{\underline{\underline{\Sigma}}}_1^{-t}(m,\zeta) \\
&- \left[\hat{\underline{\underline{T}}}_1(\zeta,m)\hat{\underline{\underline{\rho}}}_2 + \hat{\underline{\underline{T}}}_2(\zeta,m)\hat{\underline{\underline{\rho}}}_1\right]\hat{\underline{\underline{\Sigma}}}_2^{-t}(m,\zeta) \qquad (4.69)\\
-\frac{d}{d\zeta}\hat{\underline{\underline{\Sigma}}}_2^{+t}(\zeta,m) &= \hat{\underline{\underline{T}}}_1(\zeta,m)\hat{\underline{\underline{\Sigma}}}_2^{+}(\zeta) + \hat{\underline{\underline{T}}}_2(\zeta,m)\hat{\underline{\underline{\Sigma}}}_1^{+}(\zeta)) \\
&+ \left[\hat{\underline{\underline{T}}}_1(\zeta,w)\hat{\underline{\underline{\rho}}}_2 + \hat{\underline{\underline{T}}}_2(\zeta,m)\hat{\underline{\underline{\rho}}}_1\right]\hat{\underline{\underline{\Sigma}}}_2^{-t}(m,\zeta) \\
&+ \left[\hat{\underline{\underline{T}}}_1(\zeta,w)\hat{\underline{\underline{\rho}}}_1 - \hat{\underline{\underline{T}}}_2(\zeta,m)\hat{\underline{\underline{\rho}}}_2\right]\hat{\underline{\underline{\Sigma}}}_1^{-t}(m,\zeta) \qquad (4.70)
\end{aligned}
$$

The initial conditions for the upward sweep are

$$\hat{\underline{\underline{R}}}_1(m,m) = \hat{\underline{\underline{R}}}_2(m,m) = \underline{\underline{0}}_{4M\times 4M} \qquad (4.71)$$

$$\hat{\underline{\underline{T}}}_1(m,m) = \hat{\underline{\underline{T}}}_2(m,m) = \underline{\underline{I}}_{4M\times 4M} \qquad (4.72)$$

$$\hat{\underline{\underline{\Sigma}}}_{1,2}^{\pm t}(m,m) = \underline{\underline{0}}_{4M\times 1} \qquad (4.73)$$

Just as for the "bare" slab surface boundary conditions of Eqs. (4.55–4.57), these initial conditions are general for a "bare" or fully transparent bottom. However, unlike for the surface boundary, Eqs. (4.71) and (4.72) will be revised in §5.2 to incorporate the actual bottom reflectance and transmittance as the initial condition for the upward integration sweep.

4.5.3 Checks on the ODEs

I am unaware of any papers giving the preceding derivation of the ODEs for the Fourier amplitudes of the standard reflectances and transmittances for the VRTE. It is therefore desirable to have as many checks as possible on the correctness of the derivation.

A weak check for typos can be made from the Gestalt of the various equations and the patterns of subscripts, superscripts, and signs.

If the phase matrix is zero except for the (1,1) element, the VRTE reduces to the SRTE. The Fourier expansion of the phase function then contains only cosine amplitudes, which means that all of the $\hat{\underline{\underline{\rho}}}_2$, $\hat{\underline{\underline{\tau}}}_2$, $\hat{\underline{\underline{R}}}_2$ and $\hat{\underline{\underline{T}}}_2$ terms are zero. In that case Eqs. (4.40)-(4.69) all reduce to the corresponding equations (8.74)-(8.85) seen in *Light and Water* for the case of no polarization. Those equations have proven to be correct and numerically efficient via their use in HydroLight. Unfortunately, this comparison does not help to verify the correctness of the newly derived equations for the various sine amplitudes. (Strictly speaking, the VRTE matrix equations reduce to the transposes of the matrix equations seen in *Light and Water*. This is because L&W started by writing the path function as $\int L\,\tilde{\beta}\,d\Omega$, rather than as $\int \tilde{\beta}\,L\,d\Omega$ corresponding to the vector version $\int \underline{\tilde{P}}\,\underline{S}\,d\Omega$ as seen in Eq. (1.22). This led to the scalar equivalent of Eq. (4.23) being written as a row vector, and the scalar versions of $\hat{\underline{\underline{\rho}}}_1$ being the transpose of Eq. (4.24), etc.)

Ultimately, though, the only real check on the correctness of the ODEs for the VRTE standard operators is whether they give correct predictions of the Stokes vectors when numerically solved as part of the HydroPol code.

4.5.4 Computational Issues

Recalling that the 8 standard matrices are each $4M \times 4M$ in size, and the 4 internal source vectors are $4M \times 1$, this gives a set of $N_{\text{eq}} = 8 \times 4M \times 4M + 4 \times 4M = 128M^2 + 16M = 12,960$ scalar ODEs for a quad discretization with $M = 10$. The unpolarized case has no sine terms, which reduces the number of standard operators equations from 8 to 4, with 2 source-term equations, and each standard operator matrix is $M \times M$ since there is only a phase function (or (1,1) element of the phase matrix). The total number of scalar ODEs to be solved is then just $N_{\text{eq}} = 4 \times M \times M + 2 \times M = 420$. Thus there are roughly 30 times as many scalar ODEs to be solved for the VRTE as for the SRTE, namely twice as many Fourier coefficients and 4×4 phase matrices matrices rather than scalar phase functions. In both cases these sets of ODEs must be solved for each Fourier mode, $\ell = 0, ..., N$.

Many types of ODE solution algorithms exist to solve systems of equations of the form

$$\begin{aligned} \frac{d}{d\zeta} y_1(\zeta) &= f_1(\zeta, y_1, ..., y_{N_{\text{eq}}}) \\ \vdots \quad &= \quad \vdots \\ \frac{d}{d\zeta} y_{N_{\text{eq}}}(\zeta) &= f_{N_{\text{eq}}}(\zeta, y_1, ..., y_{N_{\text{eq}}}) \end{aligned}$$

These routines require as input vectors containing the values of the variables $y_i(\zeta)$ and their derivatives $d_i(\zeta)$ at the current depth ζ. The routine then returns the solution $y_i(\zeta_{\text{out}})$ at the requested output depth ζ_{out}. It is thus necessary to "unstack" the composite matrix equations to make long vectors of scalar equations.

The matrix multiplications seen in Eqs. (4.40)-(4.70) are conveniently performed using the Fortran 95 matrix multiplication intrinsic procedure MATMUL to obtain the the matrices of derivatives $\underline{D} = d\underline{Y}/d\zeta$, where $\underline{Y}$ is any of $\underline{\underline{\hat{R}}}_1(w,\zeta)$, $\underline{\underline{\hat{\Sigma}}}_1^{-t}(\zeta,w)$, etc. For a given value of ℓ and for the downward integration sweep, the (row, column) $= (I,J)$ elements of the matrix ODEs are then reformed column by column into a single vector $\underline{D} = [D_1, ..., D_{N_{\rm eq}}]$ as follows:

$$
\begin{aligned}
&\text{Jloop : DO J} = 1, 4\text{M} \\
&\quad \text{Iloop : DO I} = 1, 4\text{M} \\
&\qquad D(J + (I-1)4M + 0(4M)^2) = [\frac{d}{d\zeta}\underline{\underline{\hat{R}}}_1(\zeta, w, \ell)]_{I,J} \\
&\qquad D(J + (I-1)4M + 1(4M)^2) = [\frac{d}{d\zeta}\underline{\underline{\hat{R}}}_2(\zeta, w, \ell)]_{I,J} \\
&\qquad D(J + (I-1)4M + 2(4M)^2) = [\frac{d}{d\zeta}\underline{\underline{\hat{T}}}_1(w, \zeta, \ell)]_{I,J} \\
&\qquad D(J + (I-1)4M + 3(4M)^2) = [\frac{d}{d\zeta}\underline{\underline{\hat{T}}}_2(w, \zeta, \ell)]_{I,J} \\
&\qquad D(J + (I-1)4M = 4(4M)^2) = [\frac{d}{d\zeta}\underline{\underline{\hat{R}}}_1(w, \zeta, \ell)]_{I,J} \\
&\qquad D(J + (I-1)4M + 5(4M)^2) = [\frac{d}{d\zeta}\underline{\underline{\hat{R}}}_2(w, \zeta, \ell)]_{I,J} \\
&\qquad D(J + (I-1)4M + 6(4M)^2) = [\frac{d}{d\zeta}\underline{\underline{\hat{T}}}_1(\zeta, w, \ell)]_{I,J} \\
&\qquad D(J + (I-1)4M + 7(4M)^2) = [\frac{d}{d\zeta}\underline{\underline{\hat{T}}}_2(\zeta, w, \ell)]_{I,J} \\
&\quad \text{END DO Iloop} \\
&\qquad D(J + 0M + 8(4M)^2) = [\underline{\underline{\hat{\Sigma}}}_1^{+t}(w, \zeta, \ell)]_J \\
&\qquad D(J + 4M + 8(4M)^2) = [\underline{\underline{\hat{\Sigma}}}_2^{+t}(w, \zeta, \ell)]_J \\
&\qquad D(J + 8M + 8(4M)^2) = [\underline{\underline{\hat{\Sigma}}}_1^{-t}(\zeta, w, \ell)]_J \\
&\qquad D(J + 12M + 8(4M)^2) = [\underline{\underline{\hat{\Sigma}}}_2^{-t}(\zeta, w, \ell)]_J \\
&\text{END DO Jloop}
\end{aligned}
\tag{4.74}
$$

The same reordering of $\underline{\underline{\hat{R}}}_1(\zeta, w|\ell)$ etc. is used to create a vector $\underline{Y}$ of $\underline{\underline{R}}$, $\underline{\underline{T}}$, and $\underline{\underline{\Sigma}}$ values.

The indexing in Eq. (4.74) is written to emphasize the structure of the 1D vectors. In practice, many quantities can be pre-computed. In particular, it should be remembered from Fig. 1.5 that the phase matrix elements have zero sine amplitudes for the upper left and lower right 2×2 blocks, and zero cosine elements for the upper right and lower left 2×2 blocks. These zeros carry through the evaluation of the $\underline{\underline{\hat{\rho}}}_{1,2}$ and $\underline{\underline{\hat{\tau}}}_{1,2}$ matrices, which reduces some of the computations needed to evaluate the right hand sides of the ODEs for the standard operators. However, the ODEs for $\underline{\underline{\hat{R}}}_1$ etc. involve both cosine and sine components of the $\underline{\underline{\hat{\rho}}}_{1,2}$ and $\underline{\underline{\hat{\tau}}}_{1,2}$ matrices, so that all elements of $\underline{\underline{\hat{R}}}_1$ etc. will in general be non-zero.

If source terms are not included in the run, then the equations for $\hat{\underline{\underline{\Sigma}}}_{1,2}^{\pm t}$ can be omitted, resulting in a slightly smaller set of ODEs to be solved. A similar unwrapping of composite matrices is used for the upward sweep equations. After returning from the ODE solver, the solution vector $\underline{Y}$ is reformed into composite matrices for subsequent matrix calculations.

When running HydroPol, one of the user inputs is a list of physical depths $z_1 = w = 0$, z_2, ..., z_k, ..., $z_K = m$ at which the computed Stokes vectors are to be saved. When the downward and upward sets of ODEs are solved, the reflectance, transmittance, and source matrices are saved at these output depths for later use. That is, the arrays $\hat{\underline{\underline{R}}}_1(w, z_1{=}w{=}0|\ell) = \underline{\underline{0}}, \hat{\underline{\underline{R}}}_1(w, z_2|\ell), ..., \hat{\underline{\underline{R}}}_1(w, z_K = m|\ell)$, etc., are saved as outputs from the ODE solutions. Keep in mind that these ODE systems are solved separately for each Fourier mode $\ell = 0, 1, ...N$.

Now that we have seen the laborious development of the ODEs for the Fourier amplitudes of the standard matrices, it is worthwhile to consider the corresponding development had we stayed in physical space. The upward and downward pair of of Eqs. (4.2) would be recast into local interaction equations of the form (ignoring internal sources)

$$\frac{d}{d\zeta}\underline{\underline{S}}^{+} = \underline{\underline{\tau}}\ \underline{\underline{S}}^{+} + \underline{\underline{\rho}}\ \underline{\underline{S}}^{-}$$
$$-\frac{d}{d\zeta}\underline{\underline{S}}^{-} = \underline{\underline{\rho}}\ \underline{\underline{S}}^{+} + \underline{\underline{\tau}}\ \underline{\underline{S}}^{-}$$

Here $\underline{\underline{S}}^{\pm}$ are each $8MN \times 1$ composite vectors of 4×1 Stokes vectors $\underline{S}^{\pm}(\zeta, u, v)$ for $u = 1, ..., M$ and $v = 1, ..., 2N$. $\underline{\underline{\rho}}$ and $\underline{\underline{\tau}}$ are consequently of size $8MN \times 8MN$. However, recall that the phase matrix depends on $\phi - \phi'$, i.e. on $\phi_v - \phi_s$, so that we can set $\phi' = 0$ or $s = 1$. This reduces the number of unique elements of the $\underline{\underline{\rho}}$ and $\underline{\underline{\tau}}$ matrices by a factor of N. Thus the two $\underline{\underline{\rho}}$ and $\underline{\underline{\tau}}$ matrices each have an "information content" of $64M^2N$ independent numbers. The local interaction equations would lead to global interaction equations with four standard matrices: $\underline{\underline{R}}(w, \zeta), \underline{\underline{R}}(\zeta, w), \underline{\underline{T}}(w, \zeta)$, and $\underline{\underline{T}}(\zeta, w)$, with a total information content of $128M^2N$. It would be necessary to solve for the elements of these standard matrices via a coupled system of $128M^2N$ ODEs. For the HydroPol quad partioning with $M = 10, N = 12$ (again ignoring the internal source terms), this is a system of 153,600 equations.

Taking the Fourier-space route, we obtained global interaction equations with 4 matrices for Fourier amplitudes, $\hat{\underline{\underline{R}}}_1(w, \zeta|\ell), \hat{\underline{\underline{R}}}_1(\zeta, w|\ell), \hat{\underline{\underline{T}}}_1(w, \zeta|\ell)$, and $\hat{\underline{\underline{T}}}_1(\zeta, w|\ell)$, with each matrix being $4M \times 4M$ and $\ell = 0, ..., N$. The information content of these cosine-amplitude equations is thus $64M^2(N+1)$. There are 4 matrices for sine amplitudes, $\hat{\underline{\underline{R}}}_2(w, \zeta|\ell), \hat{\underline{\underline{R}}}_2(\zeta, w|\ell), \hat{\underline{\underline{T}}}_2(w, \zeta|\ell)$, and $\hat{\underline{\underline{T}}}_2(\zeta, w|\ell)$, with each matrix being $4M \times 4M$ and $\ell = 1, ..., N-1$. The information content of these sine-amplitude equations is thus $64M^2(N-1)$. The total number of unique matrix elements being computed is then $128M^2N$, the same as for the physical-space formulation. However, the Fourier approach allows this total to be obtained by a sequence of solutions of much smaller sets of ODEs[5].

[5]Even with the Fourier decomposition, the ODE sets are still about 13,000 equations vs 420 for the SRTE. I don't know how much of a numerical problem that will be. The physical reflectance and transmittance matrices are well behaved and bounded by 0 and 1, so presumably their Fourier amplitudes are also well behaved and smoothly varying with depth. In HydroLight I use a public (from the US National Institute of Standards and Technology) Adams-Moulton predictor-corrector ODE solver, which seems to work quite well. I have no idea how well it will work

for a set of equations 30 times larger. However, a numerical analyst tells me that the run time for ODE solvers is usually directly proportional to the number of equations being solved. Thus the run time for the vector code should be roughly 30 times that for the scalar (HydroLight) code, which is acceptable. Any recommendations on what ODE solver would be best for my system of equations are greatly appreciated.

CHAPTER 5

Incorporation of the Boundary Conditions

The matrices $\underline{\underline{\hat{R}}}_1(w, z_k|\ell)$ etc. computed from the ODE solutions give the reflectances and transmittances for subslabs $[w, z_k]$ and $[z_k, m]$ of the "bare-slab" water column $[w, m]$. These will be used to compute the Stokes vectors at the user-defined output depths z_k. However, to obtain the Stokes vectors, it is necessary to incorporate the surface and bottom boundary conditions for the particular incident lighting, surface wave conditions, and bottom reflectance of interest. Thus we must now merge the air-water surface $[a, w]$ with the water body $[w, m]$.

5.1 Fourier Decomposition of the Surface Boundary Conditions

Recall the discretized surface boundary conditions of Eqs. (2.21) and(2.22). Using the $\pm$ superscripts to denote downwelling and upwelling directions as in Eq. (4.2), the surface boundary conditions can be written

$$\underline{S}^-(a,u,v) = \sum_r \sum_s \underline{r}_{aw}(r, s \to u, v)\, \underline{S}^+(a,r,s) + \sum_r \sum_s \underline{t}_{wa}(r, s \to u, v)\, \underline{S}^-(w,r,s) \qquad (5.1)$$

$$\underline{S}^+(w,u,v) = \sum_r \sum_s \underline{r}_{wa}(r, s \to u, v)\, \underline{S}^-(w,r,s) + \sum_r \sum_s \underline{t}_{aw}(r, s \to u, v)\, \underline{S}^+(a,r,s) \qquad (5.2)$$

where the wavelength index has been omitted for brevity. The superscripts on $\underline{S}^{\pm}$ indicate whether u, v and r, s are in the upward or downward hemispheres. The 4×4 matrices $\underline{r}_{aw}(r, s \to u, v)$ etc. play the same role for surface scattering as does the phase matrix for scattering in the interior of the water body. In particular, $\underline{r}_{aw}(r, s \to u, v)$ etc. include the effects of the rotation matrices as was seen in §3.2.

As with the VRTE, the boundary conditions are Fourier decomposed. However, the surface transfer functions $\underline{r}_{aw}(r, s \to u, v)$ etc. do not in general have the $\phi - \phi' = \phi_v - \phi_s$ dependence on azimuthal angle that the phase matrix has. It is therefore necessary to do a full Fourier decomposition in ϕ_s and ϕ_v as shown in Eq. (A.15). Thus, for given r and u values, $\underline{r}_{aw}(r, s \to u, v)$

becomes

$$\begin{aligned}\underline{r}_{aw}(r,s\to u,v) = &\sum_{k=0}^{N}\sum_{\ell=0}^{N}\hat{\underline{r}}_{11}(a,w|r,u|k,\ell)\cos(k\phi_s)\cos(\ell\phi_v)\\ &+\sum_{k=0}^{N}\sum_{\ell=0}^{N}\hat{\underline{r}}_{12}(a,w|r,u|k,\ell)\cos(k\phi_s)\sin(\ell\phi_v)\\ &+\sum_{k=0}^{N}\sum_{\ell=0}^{N}\hat{\underline{r}}_{21}(a,w|r,u|k,\ell)\sin(k\phi_s)\cos(\ell\phi_v)\\ &+\sum_{k=0}^{N}\sum_{\ell=0}^{N}\hat{\underline{r}}_{22}(a,w|r,u|k,\ell)\sin(k\phi_s)\sin(\ell\phi_v)\,. \end{aligned}\tag{5.3}$$

To find $\hat{\underline{r}}_{11}$, multiple Eq. (5.3) by $\cos(k'\phi_s)\cos(\ell'\phi_v)$, sum over s and v, and apply the orthogonality relations. The other three amplitudes are found in an analogous manner. The results are as shown in Eq. (A.16):

$$\hat{\underline{r}}_{11}(a,w|r,u|k,\ell) = \frac{1}{\epsilon_k\epsilon_\ell}\sum_{s=1}^{2N}\sum_{v=1}^{2N}\underline{r}_{aw}(r,s\to u,v)\cos(k\phi_s)\cos(\ell\phi_v)\,,\tag{5.4a}$$

$$\hat{\underline{r}}_{12}(a,w|r,u|k,\ell) = \frac{1}{\epsilon_k\gamma_\ell}\sum_{s=1}^{2N}\sum_{v=1}^{2N}\underline{r}_{aw}(r,s\to u,v)\cos(k\phi_s)\sin(\ell\phi_v)\,,\tag{5.4b}$$

$$\hat{\underline{r}}_{21}(a,w|r,u|k,\ell) = \frac{1}{\gamma_k\epsilon_\ell}\sum_{s=1}^{2N}\sum_{v=1}^{2N}\underline{r}_{aw}(r,s\to u,v)\sin(k\phi_s)\cos(\ell\phi_v)\,,\tag{5.4c}$$

$$\hat{\underline{r}}_{22}(a,w|r,u|k,\ell) = \frac{1}{\gamma_k\gamma_\ell}\sum_{s=1}^{2N}\sum_{v=1}^{2N}\underline{r}_{aw}(r,s\to u,v)\sin(k\phi_s)\sin(\ell\phi_v)\,.\tag{5.4d}$$

The sine amplitudes are zero for various special cases:

$$\begin{aligned}&\hat{\underline{r}}_{12}(a,w|r,u|k,0)=\hat{\underline{r}}_{12}(a,w|k,N)=0 \qquad \text{for k} = 0,...,\text{N}\,,\\ &\hat{\underline{r}}_{21}(a,w|r,u|0,\ell)=\hat{\underline{r}}_{21}(a,w|N,\ell)=0 \qquad \text{for } \ell = 0,...,\text{N}\,,\\ &\hat{\underline{r}}_{22}(a,w|r,u|0,0)=\hat{\underline{r}}_{22}(a,w|r,u|0,N)=\hat{\underline{r}}_{22}(a,w|r,u|N,0)=\hat{\underline{r}}_{22}(a,w|r,u|N,N)=0\,. \end{aligned}\tag{5.5}$$

These special cases allow the exclusion of any k or ℓ values in Eqs. (5.4b)-(5.4d) that would result in division by zero resulting from the γ_k and γ_ℓ factors.

This expansion allows for the modeling of any sea surface. However, wave spectra generally have symmetry about the upwind-downwind direction. That is, waves may be propagating in all directions, but with equal probability to the left and right of the wind direction. Such wave spectra generally have elliptical symmetry, with the major axis of the ellipse in the along-wind direction, and the minor axis in the cross-wind direction. Thus the surface transfer functions from, say, $\phi' = \phi_s = 30\,\text{deg}$ to $\phi = \phi_v = 110$ would be the same as for $\phi' = 330$ to $\phi = 250$, but not the same as from $\phi' = 70$ to $\phi = 150$.

If the azimuthal coordinate system is chosen so that an azimuthal direction of zero is in the downwind direction, as in HydroPol (and also in HydroLight), then the statistics of the sea surface

waves propagating in some direction ϕ are the same as for direction $2\pi - \phi$. In the HydroPol quad indexing scheme of Fig. 2.1, direction ϕ_w is symmetric about the downwind direction with ϕ_{2N+2-w}. Now consider $\underline{\hat{r}}_{12}(a,w|r,u|k,\ell)$, which with these symmetries can be written

$$\underline{\hat{r}}_{12}(a,w|r,u|k,\ell) = \frac{1}{\epsilon_k \gamma_\ell} \sum_{s=1}^{2N} \sum_{v=1}^{2N} \underline{r}_{aw}(r, 2N+2-s \rightarrow u, 2N+2-v) \cos(k\phi_s) \sin(\ell\phi_v)\,.$$

Changing the summation indices to $s' = 2N+2-s$ and $v' = 2N+2-v$ gives

$$\underline{\hat{r}}_{12}(a,w|r,u|k,\ell) = \frac{1}{\epsilon_k \gamma_\ell} \sum_{s'=2N+1}^{2} \sum_{v'=2N+1}^{2} \underline{r}_{aw}(r, s' \rightarrow u, v') \cos(k\phi_{2N+2-s'}) \sin(\ell\phi_{2N+2-v'})\,.$$

Noting once again that $\phi_{2M+2-w} = 2\pi - \phi_w$, using the evenness of the cosine and oddness of the sine, remembering that quad $w = 2N+1$ is the same as quad $w = 1$, and reordering the sums gives

$$\begin{aligned}\underline{\hat{r}}_{12}(a,w|r,u|k,\ell) &= -\frac{1}{\epsilon_k \gamma_\ell} \sum_{s'=1}^{2N} \sum_{v'=1}^{2N} \underline{r}_{aw}(r, s' \rightarrow u, v') \cos(k\phi_{s'}) \sin(\ell\phi_{v'}) \\ &= -\underline{\hat{r}}_{12}(a,w|r,u|k,\ell)\,.\end{aligned}$$

It therefore follows that

$$\underline{\hat{r}}_{12}(a,w|r,u|k,\ell) = \underline{0} \qquad \text{for } r,u = 1,...,M \text{ and } k,\ell = 0,...,N\,.$$

The same analysis gives $\underline{\hat{r}}_{21}(a,w|r,u|k,\ell) = \underline{0}$, with the same result for the other three surface transfer functions. Thus *elliptical symmetry of the surface about the wind direction eliminates two of the four terms in Eq. (5.3). This is the reason for choosing a wind-centered azimuthal coordinate system.*

A similar analysis based on the elliptical symmetry between ϕ and $\pi - \phi$ gives

$$\begin{aligned}\underline{\hat{r}}_{11}(a,w|r,u|k,\ell) &= (-1)^{k+\ell} \frac{1}{\epsilon_k \epsilon_\ell} \sum_{s'=1}^{2N} \sum_{v'=1}^{2N} \underline{r}_{aw}(r, s' \rightarrow u, v') \cos(k\phi_{s'}) \sin(\ell\phi_{v'}) \\ &= (-1)^{k+\ell}\, \underline{\hat{r}}_{11}(a,w|k,\ell)\,.\end{aligned}$$

Therefor it follows that

$$\underline{\hat{r}}_{11}(a,w|r,u|k,\ell) = \underline{0} \qquad \text{if } k+\ell \text{ is odd,}$$

for $r,u = 1,...,M$ and $k,\ell = 0,...,N$. Corresponding results are obtained for $\underline{\hat{r}}_{22}(a,w|r,u|k,\ell)$ and the other three surface transfer functions.

We can therefor simplify the notation to one subscript for the $\underline{\hat{r}}$ and $\underline{\hat{t}}$ terms:

$$\begin{aligned}\underline{\hat{r}}_1(a,w|r,u|k,\ell) &= \underline{\hat{r}}_{11}(a,w|r,u|k,\ell) \\ \underline{\hat{r}}_2(a,w|r,u|k,\ell) &= \underline{\hat{r}}_{22}(a,w|r,u|k,\ell)\end{aligned}$$

since the $\underline{\hat{r}}_{12}$ and $\underline{\hat{r}}_{21}$ terms are always zero. The surface transfer functions for elliptically symmetric sea surfaces thus can be decomposed as

$$\underline{r}_{aw}(r, s \to u, v) = \sum_{\substack{k=0 \\ (k+\ell \text{ even})}}^{N} \sum_{\ell=0}^{N} \underline{\hat{r}}_1(a, w|r, u|k, \ell) \cos(k\phi_s) \cos(\ell\phi_v) + \sum_{\substack{k=0 \\ (k+\ell \text{ even})}}^{N} \sum_{\ell=0}^{N} \underline{\hat{r}}_2(a, w|r, u|k, \ell) \sin(k\phi_s) \sin(\ell\phi_v) \,. \tag{5.6}$$

where

$$\underline{\hat{r}}_1(a, w|r, u|k, \ell) = \begin{cases} \frac{1}{\epsilon_k \epsilon_\ell} \sum_{s=1}^{2N} \sum_{v=1}^{2N} \underline{r}_{aw}(r, s \to u, v) \cos(k\phi_s) \cos(\ell\phi_v) \\ \qquad \text{for } k, \ell = 0, ..., N \text{ and } k + \ell \text{ even} \\ \underline{0}_{4\times 4} \text{ if } k + \ell \text{ is odd} \end{cases} \tag{5.7}$$

and

$$\underline{\hat{r}}_2(a, w|r, u|k, \ell) = \begin{cases} \frac{1}{\gamma_k \gamma_\ell} \sum_{s=1}^{2N} \sum_{v=1}^{2N} \underline{r}_{aw}(r, s \to u, v) \cos(k\phi_s) \cos(\ell\phi_v) \\ \qquad \text{for } k, \ell = 1, ..., N-1 \text{ and } k + \ell \text{ even} \\ \underline{0}_{4\times 4} \text{ if } k = 0 \text{ and } \ell = 0, ..., N \\ \underline{0}_{4\times 4} \text{ if } \ell = 0 \text{ and } k = 0, ..., N \\ \underline{0}_{4\times 4} \text{ if } k + \ell \text{ is odd} \end{cases} \tag{5.8}$$

Equations (4.3) and (5.6) can now be used to Fourier decompose the surface boundary conditions. The process is the same as before, using orthogonality relations to reduce double sums to single, and noting the linear independence of sines and cosines. The result for the cosine amplitudes of Eq. (5.1) is

$$\underline{S}_1^-(a, u|\ell) = \sum_{\substack{k=0 \\ (k+\ell \text{ even})}}^{N} \left[\sum_{r=1}^{M-1} \epsilon_k \, \underline{\hat{r}}_1(a, w|r, u|k, \ell) \underline{S}_1^+(a, r|k) + \delta_k \, \underline{\hat{r}}_1(a, w|M, u|k, \ell) \, \underline{S}_1^+(a, M|k) \right] + \sum_{\substack{k=0 \\ (k+\ell \text{ even})}}^{N} \left[\sum_{r=1}^{M-1} \epsilon_k \, \underline{\hat{t}}_1(w, a|r, u|k, \ell) \, \underline{S}_1^-(w, r|k) + \delta_k \, \underline{\hat{t}}_1(w, a|M, u|k, \ell) \, \underline{S}_1^-(w, M|k) \right] \tag{5.9}$$

which holds for $u = 1, ..., M$ and $\ell = 0, ..., N$. As always, the polar cap $r = M$ is a special case with no ϕ dependence, so that only the $k = 0$ cosine amplitude is nonzero. However, the polar cap terms can be incorporated into the sum over k by use of a Kronecker δ_k factor, as shown here. The sine amplitudes are given by

$$\underline{S}_2^-(a, u|\ell) = \sum_{\substack{k=0 \\ (k+\ell \text{ even})}}^{N} \left[\sum_{r=1}^{M-1} \gamma_k \, \underline{\hat{r}}_2(a, w|r, u|k, \ell) \, \underline{S}_2^+(a, r|k) \right] + \sum_{\substack{k=0 \\ (k+\ell \text{ even})}}^{N} \left[\sum_{r=1}^{M-1} \gamma_k \, \underline{\hat{t}}_2(w, a|r, u|k, \ell) \, \underline{S}_2^-(w, r|k) \right] \tag{5.10}$$

where now $\ell = 1, ..., N$. As always, sine amplitudes are zero for polar caps. A similar result holds for Eq. (5.2).

It is important to note that *the Fourier modes do not decouple for the surface transfer functions.* That is, the equations for $\underline{S}^{-}_{1,2}(a, u|\ell)$ involve sums over k. This is the reason that the surface boundary conditions were not used as the initial conditions for mode-by-mode integration of the ODEs to obtain the standard reflection and transmission matrices for the interior of the water body.

These equations can be placed in a more compact form by defining $4M \times 4M$ composite matrices $\underline{\hat{\underline{r}}}_{1,2}$ and $\underline{\hat{\underline{t}}}_{1,2}$ with (row, column) $= u, r$ elements as follows:

$$[\,\underline{\hat{\underline{r}}}_1(a, w|k, \ell)]_{u,r} = \begin{cases} \epsilon_k\ \underline{\hat{r}}_1(a, w|r, u|k, \ell) & \text{for r} = 1, ..., \text{M} - 1 \\ \delta_k\ \underline{\hat{r}}_1(a, w|M, u|k, \ell) & \text{for r} = \text{M} \end{cases} \tag{5.11}$$

which holds for $u = 1, ..., M$ and $k, \ell = 0, ..., N$. The corresponding definition for the sine term is

$$[\,\underline{\hat{\underline{r}}}_2(a, w|k, \ell)]_{u,r} = \begin{cases} \gamma_k\ \underline{\hat{r}}_2(a, w|r, u|k, \ell) & \text{for } r = 1, ..., M - 1 \text{ and } \ell = 1, ..., N - 1 \\ \underline{0}_{4\times 4} & \text{for r} = \text{M or } \ell = 0 \text{ or N} \end{cases} \tag{5.12}$$

which holds for which holds for $u = 1, ..., M$ and $k = 0, ..., N$. The $\underline{\hat{t}}_1$ and $\underline{\hat{t}}_2$ functions have corresponding definitions.

Recalling the composite matrix form (4.23) for Stokes vectors, Eqs. (5.9) and (5.10) can now be combined and written as

$$\underline{\hat{\underline{S}}}^{-}_p(a|\ell) = \sum_{\substack{k=0 \\ (k+\ell \text{ even})}}^{N} \underline{\hat{\underline{r}}}_p(a, w|k, \ell)\ \underline{\hat{\underline{S}}}^{+}_p(a|\ell) + \sum_{\substack{k=0 \\ (k+\ell \text{ even})}}^{N} \underline{\hat{\underline{t}}}_p(w, a|k, \ell)\ \underline{\hat{\underline{S}}}^{-}_p(w|\ell)\,. \tag{5.13}$$

Equation (5.2) similarly becomes

$$\underline{\hat{\underline{S}}}^{+}_p(w|\ell) = \sum_{\substack{k=0 \\ (k+\ell \text{ even})}}^{N} \underline{\hat{\underline{r}}}_p(w, a|k, \ell)\ \underline{\hat{\underline{S}}}^{-}_p(w|\ell) + \sum_{\substack{k=0 \\ (k+\ell \text{ even})}}^{N} \underline{\hat{\underline{t}}}_p(a, w|k, \ell)\ \underline{\hat{\underline{S}}}^{+}_p(a|\ell)\,, \tag{5.14}$$

where $p = 1$ or 2 and $\ell = 0, ..., N$.

Equations (5.13) and (5.14) are the desired Fourier-amplitude forms of the surface boundary conditions. These equations are at a notational level equivalent to the local interaction equations (4.25) to (4.28).

Equations (5.13) and (5.14) can be notationally simplified still further by defining $4M(N+1) \times 4M(N+1)$ composite matrices that contain all Fourier modes. The $4M \times 4M$ cosine amplitude

matrices $\underline{\underline{\hat{r}}}_1(a,w|k,\ell)$ for $k,\ell = 0,1,...,N$ can be combined into one matrix as follows (L&W 8.86):

$$\underline{\underline{\hat{r}}}_1(a,w) = \begin{bmatrix} \underline{\underline{\hat{r}}}_1(a,w|0,0) & \underline{\underline{0}} & \underline{\underline{\hat{r}}}_1(a,w|2,0) & \underline{\underline{0}} & \dots & \underline{\underline{\hat{r}}}_1(a,w|N,0) \\ \underline{\underline{0}} & \underline{\underline{\hat{r}}}_1(a,w|1,1) & \underline{\underline{0}} & \underline{\underline{\hat{r}}}_1(a,w|3,1) & \dots & \underline{\underline{0}} \\ \underline{\underline{\hat{r}}}_1(a,w|0,2) & \underline{\underline{0}} & \underline{\underline{\hat{r}}}_1(a,w|2,2) & \underline{\underline{0}} & \dots & \underline{\underline{\hat{r}}}_1(a,w|N,2) \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ \underline{\underline{0}} & \underline{\underline{\hat{r}}}_1(a,w|1,N-1) & \underline{\underline{0}} & \underline{\underline{\hat{r}}}_1(a,w|3,N-1) & \dots & \underline{\underline{0}} \\ \underline{\underline{\hat{r}}}_1(a,w|0,N) & \underline{\underline{0}} & \underline{\underline{\hat{r}}}_1(a,w|2,N) & \underline{\underline{0}} & \dots & \underline{\underline{\hat{r}}}_1(a,w|N,N) \end{bmatrix} \tag{5.15}$$

The checkerboard structure of this matrix incorporates the previous result that $\underline{\underline{\hat{r}}}_1(a,w|k,\ell) \neq \underline{\underline{0}}$ only when $k+\ell$ is even. Recall than N is always even. The $\underline{\underline{0}}$ submatrices are all $4M \times 4M$. Similarly, the sine amplitude matrices $\underline{\underline{\hat{r}}}_2(a,w|k,\ell)$ can be combined as

$$\underline{\underline{\hat{r}}}_2(a,w) = \begin{bmatrix} \underline{\underline{0}} & \underline{\underline{0}} & \underline{\underline{0}} & \underline{\underline{0}} & \dots & \underline{\underline{0}} \\ \underline{\underline{0}} & \underline{\underline{\hat{r}}}_2(a,w|1,1) & \underline{\underline{0}} & \underline{\underline{\hat{r}}}_2(a,w|3,1) & \dots & \underline{\underline{0}} \\ \underline{\underline{0}} & \underline{\underline{0}} & \underline{\underline{\hat{r}}}_2(a,w|2,2) & \underline{\underline{0}} & \dots & \underline{\underline{0}}) \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ \underline{\underline{0}} & \underline{\underline{\hat{r}}}_2(a,w|1,N-1) & \underline{\underline{0}} & \underline{\underline{\hat{r}}}_2(a,w|3,N-1) & \dots & \underline{\underline{0}} \\ \underline{\underline{0}} & \underline{\underline{0}} & \underline{\underline{0}} & \underline{\underline{0}} & \dots & \underline{\underline{0}} \end{bmatrix}. \tag{5.16}$$

The $\underline{\underline{0}}$ first and last row and column show that sine amplitudes are zero for $k,\ell = 0$ or N. The $\underline{\underline{\hat{t}}}_{1,2}$ matrices have corresponding definitions. The Stokes vector amplitudes $\underline{\hat{S}}^{\pm}_{1,2}(\zeta|\ell)$ are likewise combined into one $4M(N+1) \times 1$ column vector containing all Fourier modes (L&W 8.87):

$$\underline{\hat{S}}^{\pm}_{1,2}(\zeta) = \begin{bmatrix} \underline{\hat{S}}^{\pm}_{1,2}(\zeta|0) \\ \underline{\hat{S}}^{\pm}_{1,2}(\zeta|1) \\ \vdots \\ \underline{\hat{S}}^{\pm}_{1,2}(\zeta|N) \end{bmatrix}. \tag{5.17}$$

It is worth pausing to note that the $\underline{\hat{S}}^{\pm}_{1,2}(\zeta)$ are matrices whose elements $\underline{\hat{S}}^{\pm}_{1,2}(\zeta|\ell)$ are by Eq. (4.23) $4M \times 1$ matrices, whose elements in turn are 4×1 Stokes vectors. In other words, we are now dealing with matrices whose elements are matrices whose elements are matrices.

Equations (5.15) - (5.17) allow the surface boundary conditions (5.13) and (5.14) to be written as matrix equations:

$$\underline{\hat{S}}^{-}_{p}(a) = \underline{\underline{\hat{r}}}_p(a,w)\underline{\hat{S}}^{+}_{p}(a) + \underline{\underline{\hat{t}}}_p(w,a)\underline{\hat{S}}^{-}_{p}(w) \tag{5.18}$$

and

$$\underline{\hat{S}}^{+}_{p}(w) = \underline{\underline{\hat{r}}}_p(w,a)\underline{\hat{S}}^{-}_{p}(w) + \underline{\underline{\hat{t}}}_p(a,w)\underline{\hat{S}}^{+}_{p}(a) \tag{5.19}$$

It should be noted that $\underline{\underline{\hat{r}}}_{1,2}$ and $\underline{\underline{\hat{t}}}_{1,2}$ are in effect the standard reflection and transmission matrices for the air-water surface. Unlike the standard matrices for the interior of the water body

$[w,\zeta]$, which must be found by depth integrations of the ODEs derived in the preceding chapter, the surface $[a,w]$ has no thickness and standard matrices are determined from the reflection and transmission matrices for the sea surface, as seen in Chapter 3.

Finally, the standard reflectance and transmittance matrices for the water body $[w,\zeta]$, which were computed ℓ mode by ℓ mode when solving the ODE systems, can be combined into matrices containing all Fourier modes, e.g.:

$$\underline{\underline{\hat{R}}}_1(\zeta,w) = \begin{bmatrix} \underline{\underline{\hat{R}}}_1(\zeta,w|0) & \underline{\underline{0}} & \underline{\underline{0}} & \dots & \underline{\underline{0}} \\ \underline{\underline{0}} & \underline{\underline{\hat{R}}}_1(\zeta,w|1)) & \underline{\underline{0}} & \dots & \underline{\underline{0}} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \underline{\underline{0}} & \underline{\underline{0}} & \underline{\underline{0}} & \dots & \underline{\underline{\hat{R}}}_1(\zeta,w|N) \end{bmatrix} \tag{5.20}$$

with corresponding definitions for the other standard matrices.

These definitions allow the global interaction equations (4.29)-(4.32) for slab $[w,\zeta]$ to be written in a form that contains all Fourier modes:

$$\underline{\hat{S}}_1^-(w) = \underline{\underline{\hat{R}}}_1(w,\zeta)\underline{\hat{S}}_1^+(w) + \underline{\underline{\hat{T}}}_1(\zeta,w)\underline{\hat{S}}_1^-(\zeta) - \underline{\underline{\hat{R}}}_2(w,\zeta)\underline{\hat{S}}_2^+(w) - \underline{\underline{\hat{T}}}_2(\zeta,w)\underline{\hat{S}}_2^-(\zeta) + \underline{\hat{\Sigma}}_1^{-t}(\zeta,w) \tag{5.21}$$

$$\underline{\hat{S}}_1^+(\zeta) = \underline{\underline{\hat{R}}}_1(\zeta,w)\underline{\hat{S}}_1^-(\zeta) + \underline{\underline{\hat{T}}}_1(w,\zeta)\underline{\hat{S}}_1^+(w) - \underline{\underline{\hat{R}}}_2(\zeta,w)\underline{\hat{S}}_2^-(\zeta) - \underline{\underline{\hat{T}}}_2(w,\zeta)\underline{\hat{S}}_2^+(w) + \underline{\hat{\Sigma}}_1^{+t}(w,\zeta) \tag{5.22}$$

$$\underline{\hat{S}}_2^-(w) = \underline{\underline{\hat{R}}}_1(w,\zeta)\underline{\hat{S}}_2^+(w) + \underline{\underline{\hat{T}}}_1(\zeta,w)\underline{\hat{S}}_2^-(\zeta) + \underline{\underline{\hat{R}}}_2(w,\zeta)\underline{\hat{S}}_1^+(w) + \underline{\underline{\hat{T}}}_2(\zeta,w)\underline{\hat{S}}_1^-(\zeta) + \underline{\hat{\Sigma}}_2^{-t}(\zeta,w) \tag{5.23}$$

$$\underline{\hat{S}}_2^+(\zeta) = \underline{\underline{\hat{R}}}_1(\zeta,w)\underline{\hat{S}}_2^-(\zeta) + \underline{\underline{\hat{T}}}_1(w,\zeta)\underline{\hat{S}}_2^+(w) + \underline{\underline{\hat{R}}}_2(\zeta,w)\underline{\hat{S}}_1^-(\zeta) + \underline{\underline{\hat{T}}}_2(w,\zeta)\underline{\hat{S}}_1^+(w) + \underline{\hat{\Sigma}}_2^{+t}(w,\zeta) \tag{5.24}$$

Here the internal source terms for each Fourier mode have been combined as in Eq. (5.17).

5.2 Fourier Decomposition of the Bottom Boundary Conditions

As noted in §1.6, the bottom boundary condition has the form (dropping the wavelength dependence for brevity)

$$\begin{aligned} \underline{S}(m,\theta,\phi) &= \iint_{2\pi_d} \underline{VBRDF}(\theta',\theta,\phi-\phi')\,\cos\theta'\,\underline{S}(m,\theta',\phi')\,d\Omega(\theta',\phi') \\ &= \iint_{2\pi_d} \underline{r}_{mb}(\theta',\theta,\phi-\phi')\,\underline{S}(m,\theta',\phi')\,d\Omega(\theta',\phi') \quad \text{for } (\theta,\phi)\in 2\pi_{\mathrm{u}}\ . \end{aligned}$$

A $\underline{\text{VBRDF}}$ or $\underline{r}_{mb}$ of this form can be expanded as Fourier cosine series. As was seen in Eq. (2.23), the corresponding quad-averaged bottom boundary condition is

$$\underline{S}(m,u,v) = \sum_{\substack{r \\ r,s\in\Xi_d}} \sum_{s} \underline{r}_{mb}(r,s\to u,v)\,\underline{S}(m,r,s) \quad \text{for } (\mathrm{u},\mathrm{v})\in\Xi_{\mathrm{u}}\ . \tag{5.25}$$

The Fourier decomposition of the bottom boundary condition conceptually follows in parallel to the decomposition of the surface boundary condition, but the equations are much simpler because of the $\phi-\phi'$ symmetry and the absence of transmission terms as discussed in §1.6. The Stokes vectors are expanded as in Eq. (4.3):

$$\underline{S}^{\pm}(m,u,v) = \sum_{\ell=0}^{N} \left[\underline{\hat{S}}_1^{\pm}(m,u|\ell)\cos(\ell\phi_v) + \underline{\hat{S}}_2^{\pm}(m,u|\ell)\sin(\ell\phi_v)\right]\ , \tag{5.26}$$

where as always $u = 1, 2, ..., M$ and $v = 1, 2, ..., 2N$. $\underline{r}_{mb}(r, s \to u, v) = \underline{r}_{mb}(r, u, |v - s|)$ is expanded as a cosine series (L&W 8.54):

$$\underline{r}_{mb}(r, u, |v - s|) = \sum_{k=0}^{N} \underline{\hat{r}}_1(m, b|r, u|k) \cos[k(\phi_v - \phi_s)] \,, \tag{5.27}$$

where (L&W 8.55)

$$\underline{\hat{r}}_1(m, b|r, u|k) = \frac{1}{\epsilon_k \cos(k\phi_s)} \sum_{v=1}^{2N} \underline{r}_{mb}(r, u, |v - s|) \cos(k\phi_v) \,.$$

We are again free to choose $s = 1$, or $\phi_s = 0$, which gives

$$\underline{\hat{r}}_1(m, b|r, u|k) = \frac{1}{\epsilon_k} \sum_{v=1}^{2N} \underline{r}_{mb}(r, u, |v - 1|) \cos(k\phi_v) \,. \tag{5.28}$$

Substituting these expansions into Eq. (5.25) gives

$$\begin{aligned} &\sum_{\ell=0}^{N} \left[\underline{\hat{S}}_1^-(m, u|\ell) \cos(\ell\phi_v) + \underline{\hat{S}}_2^-(m, u|\ell) \sin(\ell\phi_v) \right] \\ &= \sum_r \sum_s \sum_{k=0}^{N} \underline{\hat{r}}_1(m, b|r, u|k) \cos[k(\phi_v - \phi_s)] \left\{ \sum_{\ell=0}^{N} \left[\underline{\hat{S}}_1^+(m, r|\ell) \cos(\ell\phi_s) + \underline{\hat{S}}_2^+(m, r|\ell) \sin(\ell\phi_s) \right] \right\} \end{aligned} \tag{5.29}$$

Expanding $\cos[k(\phi_v - \phi_s)]$ and grouping terms gives

$$\begin{aligned} &\sum_{\ell=0}^{N} \left[\underline{\hat{S}}_1^-(m, u|\ell) \cos(\ell\phi_v) + \underline{\hat{S}}_2^-(m, u|\ell) \sin(\ell\phi_v) \right] \\ &= \sum_{\ell=0}^{N} \sum_r \sum_{k=1}^{N} \underline{\hat{r}}_1(m, b|r, u|k) \cos(k\phi_v) \, \underline{\hat{S}}_1^+(m, r|\ell) \sum_s \cos(k\phi_s) \cos(\ell\phi_s) \\ &= \sum_{\ell=0}^{N} \sum_r \sum_{k=1}^{N} \underline{\hat{r}}_1(m, b|r, u|k) \cos(k\phi_v) \, \underline{\hat{S}}_2^+(m, r|\ell) \sum_s \cos(k\phi_s) \sin(\ell\phi_s) \\ &= \sum_{\ell=0}^{N} \sum_r \sum_{k=1}^{N} \underline{\hat{r}}_1(m, b|r, u|k) \sin(k\phi_v) \, \underline{\hat{S}}_1^+(m, r|\ell) \sum_s \sin(k\phi_s) \cos(\ell\phi_s) \\ &= \sum_{\ell=0}^{N} \sum_r \sum_{k=1}^{N} \underline{\hat{r}}_1(m, b|r, u|k) \sin(k\phi_v) \, \underline{\hat{S}}_2^+(m, r|\ell) \sum_s \sin(k\phi_s) \sin(\ell\phi_s) \end{aligned} \tag{5.30}$$

Applying the orthogonality relations (A.3–A.5) to the sums over s and then observing as before

that $k = \ell$ follows, this reduces to

$$\begin{aligned}
&\sum_{\ell=0}^{N} \left[\hat{\underline{S}}_1^-(m,u|\ell)\cos(\ell\phi_v) + \hat{\underline{S}}_2^-(m,u|\ell)\sin(\ell\phi_v) \right] \\
&= \sum_{\ell=0}^{N} \sum_r \hat{\underline{r}}_1(m,b|r,u|k)\, \hat{\underline{S}}_1^+(m,r|\ell)\, \epsilon_\ell \cos(\ell\phi_v) \\
&+ \sum_{\ell=0}^{N} \sum_r \hat{\underline{r}}_1(m,b|r,u|k)\, \hat{\underline{S}}_2^+(m,r|\ell)\, \gamma_\ell \sin(\ell\phi_v)
\end{aligned} \tag{5.31}$$

where ϵ_ℓ and γ_ℓ are given by Eqs. (A.8) and (A.10, respectively. As always, the polar caps $u, r = M$ are special cases with no ϕ dependence.

Invoking the linear independence of $\sin(\ell\phi_v)$ and $\cos(\ell\phi_v)$ gives two separate equations relating $\hat{\underline{S}}_1^\pm(m,u|\ell)$ and $\hat{\underline{S}}_2^\pm(m,u|\ell)$. These can be put into matrix form by defining $4M \times 4M$ composite matrices with 4×4 elements (L&W 8.57):

$$[\,\hat{\underline{\underline{r}}}_1(m,b|\ell)]_{u,r} = \begin{cases} \epsilon_\ell\, \hat{\underline{r}}_1(m,b|r,u|\ell) & \text{if } r = 1,...,M-1 \\ \delta_\ell\, \hat{\underline{r}}_1(m,b|r,u|\ell) & \text{if } r = M \end{cases} \tag{5.32}$$

for $u = 1,...,M$ and $\ell = 0,...,N$; and

$$[\,\hat{\underline{\underline{r}}}_2(m,b|\ell)]_{u,r} = \begin{cases} \gamma_\ell\, \hat{\underline{r}}_2(m,b|r,u|\ell) & \text{if } r = 1,...,M-1 \text{ and } \ell = 1,...N-1 \\ 0 \quad \text{if } r = M \text{ or } \ell = 0 \text{ or } \ell = N \end{cases} \tag{5.33}$$

Equation (5.31) thus becomes

$$\begin{aligned}
\hat{\underline{S}}_1^-(m,u|\ell) &= \sum_r [\,\hat{\underline{\underline{r}}}_1(m,b|\ell)]_{u,r}\, \hat{\underline{S}}_1^+(m,r|\ell) \\
\hat{\underline{S}}_2^-(m,u|\ell) &= \sum_r [\,\hat{\underline{\underline{r}}}_2(m,b|\ell)]_{u,r}\, \hat{\underline{S}}_2^+(m,r|\ell)
\end{aligned}$$

In composite matrix form these equations are (L&W 8.56)

$$\begin{aligned}
\hat{\underline{\underline{S}}}_1^-(m|\ell) &= \hat{\underline{\underline{r}}}_1(m,b|\ell)\hat{\underline{\underline{S}}}_1^+(m|\ell) \\
\hat{\underline{\underline{S}}}_2^-(m|\ell) &= \hat{\underline{\underline{r}}}_2(m,b|\ell)\hat{\underline{\underline{S}}}_2^+(m|\ell)
\end{aligned} \tag{5.34}$$

where $\hat{\underline{\underline{S}}}_{1,2}^\pm$ are $4M \times 1$ composite column vectors defined in Eq. (4.23).

Note that, unlike for the surface boundary conditions, *the ℓ modes decouple for the bottom boundary condition because of the restriction to a bottom VBRDF dependence on* $\cos(\phi - \phi')$. *This allows the initial condition (4.71) for the upward sweep of ODEs to be revised to incorporate the bottom boundary reflectance as the initial condition for integration of the upward set of equations.* Thus Eq. (4.71),

$$\hat{\underline{\underline{R}}}_1(m,m) = \hat{\underline{\underline{R}}}_2(m,m) = \underline{\underline{0}}_{4M\times 4M}\,,$$

is replaced by (L&W 8.94)

$$\begin{aligned}
\hat{\underline{\underline{R}}}_1(m,m) &= \hat{\underline{\underline{r}}}_1(m,b|\ell) \\
\hat{\underline{\underline{R}}}_2(m,m) &= \hat{\underline{\underline{r}}}_2(m,b|\ell)
\end{aligned} \tag{5.35}$$

for each ℓ value of the upward sweep integrations. Similarly, the bottom boundary is assumed to be opaque, so that no light is transmitted upward through the bottom boundary. The initial conditions for the upward sweep integration for a reflecting, opaque, source-free lower boundary are then

$$\underline{\underline{\hat{R}}}_1(m,b) = \underline{\underline{\hat{r}}}_1(m,b|\ell) \tag{5.36}$$

$$\underline{\underline{\hat{R}}}_2(m,b) = \underline{\underline{\hat{r}}}_2(m,b|\ell) \tag{5.37}$$

$$\underline{\underline{\hat{T}}}_1(m,b) = \underline{\underline{\hat{T}}}_2(m,b) = \underline{\underline{0}}_{4M\times 4M} \tag{5.38}$$

$$\underline{\hat{\Sigma}}^{\pm t}_{1,2}(m,b) = \underline{0}_{4M\times 1} \tag{5.39}$$

Finally, recall from Eq. (1.45) that for a Lambertian depolarizing bottom

$$\underline{r}_{mb} \equiv \underline{r}_{\text{Lamb}} = \frac{R(\lambda)\cos\theta'}{\pi}\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

This implies that only the (1,1), $\ell = 0$ element of $\underline{\hat{r}}_1(m,b|\ell)$ is nonzero, and that $\underline{\hat{r}}_2(m,b|\ell) = 0$ for all ℓ values. In particular,

$$[\,\underline{\hat{r}}_1(m,b|0)]_{1,1} = \frac{R(\lambda)}{\pi}\mu_r\Omega_{r1} = \underline{r}_{mb}(r, 1 \to u, v)\,. \tag{5.40}$$

5.3 Combining the Sea Surface with the Water Body

We now have the Fourier amplitude forms of the standard reflectance and transmittance matrices for the air-water surface $[a,w]$, $\underline{\hat{r}}_1(a,w), \underline{\hat{t}}_1(a,w)$, etc., from §5.1. The standard matrices $\underline{\underline{\hat{R}}}_1(w,\zeta), \underline{\underline{\hat{T}}}_1(w,\zeta)$, etc. for the "bare slab" water body $[w,\zeta]$ are known from the solutions of the ODEs in §4.5. The next step is to combine these two slabs to obtain matrices for the air-water surface plus the water body. That is, we need to add the "real" sea surface to the top of the bare slab water body.

The goal is to obtain the response radiances at the sea surface, $\underline{\hat{S}}^{-}_{1,2}(a)$, and at depth ζ, $\underline{\hat{S}}^{+}_{1,2}(\zeta)$, in terms of the known incident radiance at the surface, $\underline{\hat{S}}^{+}_{1,2}(a)$, and at the bottom, $\underline{\hat{S}}^{-}_{1,2}(\zeta)$. This is a two-stop process:

1. The first step is to obtain the internal radiances $\underline{\hat{S}}^{\pm}_{1,2}(w)$ at depth w within the combined slab $[a,\zeta] = [a,w] \cup [w,\zeta]$ as functions of the incident radiances at a and ζ.
2. Those results will then be used to eliminate $\underline{\hat{S}}^{\pm}_{1,2}(w)$ in the surface boundary condition (5.18) and the global interaction Eqs. (5.22) and (5.24).

To obtain a general result for the union of any two slabs, we include the possibility of internal sources within $[a,w]$. When $[a,w]$ is just a discontinuity in optical properties, as for an air-water

surface, those sources are zero. However, the surface $[a, w]$ also could represent a very thin layer of oil, which can fluoresce. In that case, the surface slab $[a, w]$ would include an internal source.

Surface boundary condition (5.19) for $[a, w]$ becomes, for $p = 1, 2$ and with internal source terms added for complete generality (consider, for example, a thin surface film of fluorescing oil),

$$\hat{\underline{\underline{S}}}_1^+(w) = \hat{\underline{\underline{r}}}_1(w,a)\hat{\underline{\underline{S}}}_1^-(w) + \hat{\underline{\underline{t}}}_1(a,w)\hat{\underline{\underline{S}}}_1^+(a) + \hat{\underline{\underline{\sigma}}}_1^{+t}(a,w)$$
$$\hat{\underline{\underline{S}}}_2^+(w) = \hat{\underline{\underline{r}}}_2(w,a)\hat{\underline{\underline{S}}}_2^-(w) + \hat{\underline{\underline{t}}}_2(a,w)\hat{\underline{\underline{S}}}_2^+(a) + \hat{\underline{\underline{\sigma}}}_2^{+t}(a,w)$$

The global interaction equations (5.21) and (5.23) for $[w, \zeta]$ are

$$\hat{\underline{\underline{S}}}_1^-(w) = \hat{\underline{\underline{R}}}_1(w,\zeta)\hat{\underline{\underline{S}}}_1^+(w) + \hat{\underline{\underline{T}}}_1(\zeta,w)\hat{\underline{\underline{S}}}_1^-(\zeta) - \hat{\underline{\underline{R}}}_2(w,\zeta)\hat{\underline{\underline{S}}}_2^+(w) - \hat{\underline{\underline{T}}}_2(\zeta,w)\hat{\underline{\underline{S}}}_2^-(\zeta) + \hat{\underline{\underline{\Sigma}}}_1^{-t}(\zeta,w)$$
$$\hat{\underline{\underline{S}}}_2^-(w) = \hat{\underline{\underline{R}}}_1(w,\zeta)\hat{\underline{\underline{S}}}_2^+(w) + \hat{\underline{\underline{T}}}_1(\zeta,w)\hat{\underline{\underline{S}}}_2^-(\zeta) + \hat{\underline{\underline{R}}}_2(w,\zeta)\hat{\underline{\underline{S}}}_1^+(w) + \hat{\underline{\underline{T}}}_2(\zeta,w)\hat{\underline{\underline{S}}}_1^-(\zeta) + \hat{\underline{\underline{\Sigma}}}_2^{-t}(\zeta,w)$$

These four equations can be solved for the four internal radiances $\hat{\underline{\underline{S}}}_{1,2}^{\pm}(w)$ in terms of the four incident radiances $\hat{\underline{\underline{S}}}_{1,2}^{+}(a)$ and $\hat{\underline{\underline{S}}}_{1,2}^{-}(\zeta)$ and the given internal sources. For simplicity of notation in obtaining this solution, let $r_p = \hat{\underline{\underline{r}}}_p(w,a)$, $t_p = \hat{\underline{\underline{t}}}_p(a,w)$, $R_1 = \hat{\underline{\underline{R}}}_1(w,\zeta)$,..., $T_2 = \hat{\underline{\underline{T}}}_2(\zeta,w)$, $\sigma_p = \hat{\underline{\underline{\sigma}}}_p^{+t}(a,w)$ and $\Sigma_p = \hat{\underline{\underline{\Sigma}}}_p^{-t}(\zeta,w)$. Keep in mind that r_1 etc. are all $4M(N+1) \times 4M(N+1)$ composite matrices, and that σ_p and Σ_p are $4M(N+1) \times 1$. These equations then can be placed in matrix form as

$$\begin{bmatrix} \hat{\underline{\underline{S}}}_1^+(w) \\ \hat{\underline{\underline{S}}}_2^+(w) \\ \hat{\underline{\underline{S}}}_1^-(w) \\ \hat{\underline{\underline{S}}}_2^-(w) \end{bmatrix} = \begin{bmatrix} 0 & 0 & r_1 & 0 \\ 0 & 0 & 0 & r_2 \\ R_1 & -R_2 & 0 & 0 \\ R_2 & R_1 & 0 & 0 \end{bmatrix} \begin{bmatrix} \hat{\underline{\underline{S}}}_1^+(w) \\ \hat{\underline{\underline{S}}}_2^+(w) \\ \hat{\underline{\underline{S}}}_1^-(w) \\ \hat{\underline{\underline{S}}}_2^-(w) \end{bmatrix} + \begin{bmatrix} t_1 & 0 & 0 & 0 \\ 0 & t_2 & 0 & 0 \\ 0 & 0 & T_1 & -T_2 \\ 0 & 0 & T_2 & T_1 \end{bmatrix} \begin{bmatrix} \hat{\underline{\underline{S}}}_1^+(a) \\ \hat{\underline{\underline{S}}}_2^+(a) \\ \hat{\underline{\underline{S}}}_1^-(\zeta) \\ \hat{\underline{\underline{S}}}_2^-(\zeta) \end{bmatrix} + \begin{bmatrix} \sigma_1 \\ \sigma_2 \\ \Sigma_1 \\ \Sigma_2 \end{bmatrix}.$$

The solution is then

$$\begin{bmatrix} \hat{\underline{\underline{S}}}_1^+(w) \\ \hat{\underline{\underline{S}}}_2^+(w) \\ \hat{\underline{\underline{S}}}_1^-(w) \\ \hat{\underline{\underline{S}}}_2^-(w) \end{bmatrix} = \begin{bmatrix} I & 0 & -r_1 & 0 \\ 0 & I & 0 & -r_2 \\ -R_1 & R_2 & I & 0 \\ -R_2 & -R_1 & 0 & I \end{bmatrix}^{-1} \left\{ \begin{bmatrix} t_1 & 0 & 0 & 0 \\ 0 & t_2 & 0 & 0 \\ 0 & 0 & T_1 & -T_2 \\ 0 & 0 & T_2 & T_1 \end{bmatrix} \begin{bmatrix} \hat{\underline{\underline{S}}}_1^+(a) \\ \hat{\underline{\underline{S}}}_2^+(a) \\ \hat{\underline{\underline{S}}}_1^-(\zeta) \\ \hat{\underline{\underline{S}}}_2^-(\zeta) \end{bmatrix} + \begin{bmatrix} \sigma_1 \\ \sigma_2 \\ \Sigma_1 \\ \Sigma_2 \end{bmatrix} \right\}. \tag{5.41}$$

The needed inverse matrix can be obtained using the block matrix inversion formula

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix}^{-1} = \begin{bmatrix} (A - BD^{-1}C)^{-1} & -A^{-1}B(D - CA^{-1}B)^{-1} \\ -D^{-1}C(A - BD^{-1}C)^{-1} & (D - CA^{-1}B)^{-1} \end{bmatrix} \tag{5.42}$$

where

$$A = D = \begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix},\ B = \begin{bmatrix} -r_1 & 0 \\ 0 & -r_2 \end{bmatrix},\ \text{and } C = \begin{bmatrix} -R_1 & R_2 \\ -R_2 & -R_1 \end{bmatrix}.$$

The result is

$$\begin{bmatrix} I & B \\ C & I \end{bmatrix}^{-1} = \begin{bmatrix} (I-BC)^{-1} & -B(I-CB)^{-1} \\ -C(I-BC)^{-1} & (I-CB)^{-1} \end{bmatrix} \tag{5.43}$$

To return to the 4×4 form of Eq. (5.41), we apply Eq. (5.42) to each of the blocks of the inverse:

$$\begin{aligned}
(I-BC)^{-1} &= \begin{bmatrix} I - r_1R_1 & r_1R_2 \\ -r_2R_2 & I - r_2R_1 \end{bmatrix}^{-1} \\
&= \begin{bmatrix} [I - r_1R_1 + r_1R_2(I - r_2R_1)^{-1}r_2R_2]^{-1} \\ (I - r_2R_1)^{-1}r_2R_2[I - r_1R_1 + r_1R_2(I - r_2R_1)^{-1}r_2R_2]^{-1} \end{bmatrix} \\
&\qquad \begin{matrix} -(I - r_1R_1)^{-1}r_1R_2[I - r_2R_1 + r_2R_2(I - r_1R_1)^{-1}r_1R_2]^{-1} \\ [I - r_2R_1 + r_2R_2(I - r_1R_1)^{-1}r_1R_2]^{-1} \end{bmatrix} \equiv \begin{bmatrix} m_{11} & m_{12} \\ m_{21} & m_{22} \end{bmatrix}
\end{aligned} \tag{5.44}$$

$$-C(I-BC)^{-1} = \begin{bmatrix} R_1 & -R_2 \\ R_2 & R_1 \end{bmatrix} \begin{bmatrix} m_{11} & m_{12} \\ m_{21} & m_{22} \end{bmatrix} = \begin{bmatrix} R_1m_{11} - R_2m_{21} & R_1m_{12} - R_2m_{22} \\ R_2m_{11} + R_1m_{21} & R_2m_{12} + R_1m_{22} \end{bmatrix}$$

$$\begin{aligned}
(I-CB)^{-1} &= \begin{bmatrix} I - R_1r_1 & R_2r_2 \\ -R_2r_1 & I - R_1r_2 \end{bmatrix}^{-1} \\
&= \begin{bmatrix} [I - R_1r_1 + R_2r_2(I - R_1r_2)^{-1}R_2r_1]^{-1} \\ (I - R_1r_2)^{-1}R_2r_1[I - R_1r_1 + R_2r_2(I - R_1r_2)^{-1}R_2r_1]^{-1} \end{bmatrix} \\
&\qquad \begin{matrix} -(I - R_1r_1)^{-1}R_2r_2[I - R_1r_2 + R_2r_1(I - R_1r_1)^{-1}R_2r_2]^{-1} \\ [I - R_1r_2 + R_2r_1(I - R_1r_1)^{-1}R_2r_2]^{-1} \end{bmatrix} \equiv \begin{bmatrix} m_{33} & m_{34} \\ m_{43} & m_{44} \end{bmatrix}
\end{aligned} \tag{5.45}$$

and

$$-B(I-CB)^{-1} = \begin{bmatrix} r_1 & 0 \\ 0 & r_2 \end{bmatrix} \begin{bmatrix} m_{33} & m_{34} \\ m_{43} & m_{44} \end{bmatrix} = \begin{bmatrix} r_1m_{33} & r_1m_{34} \\ r_2m_{43} & r_2m_{44} \end{bmatrix}$$

Equation (5.41) now becomes

$$\begin{bmatrix} \hat{\underline{\underline{S}}}_1^+(w) \\ \hat{\underline{\underline{S}}}_2^+(w) \\ \hat{\underline{\underline{S}}}_1^-(w) \\ \hat{\underline{\underline{S}}}_2^-(w) \end{bmatrix} = \begin{bmatrix} m_{11}t_1 & m_{12}t_2 \\ m_{21}t_1 & m_{22}t_2 \\ (R_1m_{11} - R_2m_{21})t_1 & (R_1m_{12} - R_2m_{22})t_2 \\ (R_2m_{11} + R_1m_{21})t_1 & (R_2m_{12} + R_1m_{22})t_2 \end{bmatrix}$$

$$\begin{bmatrix} r_1m_{33}T_1 + r_1m_{34}T_2 & -r_1m_{33}T_2 + r_1m_{34}T_1 \\ r_2m_{43}T_1 + r_2m_{44}T_2 & -r_2m_{43}T_2 + r_2m_{44}T_1 \\ m_{33}T_1 + m_{34}T_2 & -m_{33}T_2 + m_{34}T_1 \\ m_{43}T_1 + m_{44}T_2 & -m_{43}T_2 + m_{44}T_1 \end{bmatrix} \begin{bmatrix} \hat{\underline{\underline{S}}}_1^+(a) \\ \hat{\underline{\underline{S}}}_2^+(a) \\ \hat{\underline{\underline{S}}}_1^-(\zeta) \\ \hat{\underline{\underline{S}}}_2^-(\zeta) \end{bmatrix}$$

$$+ \begin{bmatrix} m_{11}\sigma_1 + m_{12}\sigma_2 + r_1m_{33}\Sigma_1 + r_1m_{34}\Sigma_2 \\ m_{21}\sigma_1 + m_{22}\sigma_2 + r_2m_{43}\Sigma_1 + r_2m_{44}\Sigma_2 \\ (R_1m_{11} - R_2m_{21})\sigma_1 + (R_1m_{12} - R_2m_{22})\sigma_2 + m_{33}\Sigma_1 + m_{34}\Sigma_2 \\ (R_2m_{11} + R_1m_{21})\sigma_1 + (R_2m_{12} + R_1m_{22})\sigma_2 + m_{43}\Sigma_1 + m_{44}\Sigma_2 \end{bmatrix} \tag{5.46}$$

5.3.1 Invariant Imbedding Relations and Imbed Rules for the Surface Plus the Water Body

Following the notation of *Light and Water* (L&W 7.65-7.72 and 8.95-8.97), the matrix elements of the preceding equation are rewritten as (L&W 7.65-7.66)

$$\begin{bmatrix} \hat{\underline{\underline{S}}}_1^+(w) \\ \hat{\underline{\underline{S}}}_2^+(w) \\ \hat{\underline{\underline{S}}}_1^-(w) \\ \hat{\underline{\underline{S}}}_2^-(w) \end{bmatrix} \equiv \begin{bmatrix} \mathfrak{T}_{11}(a,w,\zeta) & \mathfrak{T}_{12}(a,w,\zeta) & \mathfrak{R}_{11}(\zeta,w,a) & \mathfrak{R}_{12}(\zeta,w,a) \\ \mathfrak{T}_{21}(a,w,\zeta) & \mathfrak{T}_{22}(a,w,\zeta) & \mathfrak{R}_{21}(\zeta,w,a) & \mathfrak{R}_{22}(\zeta,w,a) \\ \mathfrak{R}_{11}(a,w,\zeta) & \mathfrak{R}_{12}(a,w,\zeta) & \mathfrak{T}_{11}(\zeta,w,a) & \mathfrak{T}_{12}(\zeta,w,a) \\ \mathfrak{R}_{21}(a,w,\zeta) & \mathfrak{R}_{22}(a,w,\zeta) & \mathfrak{T}_{21}(\zeta,w,a) & \mathfrak{T}_{22}(\zeta,w,a) \end{bmatrix} \begin{bmatrix} \hat{\underline{\underline{S}}}_1^+(a) \\ \hat{\underline{\underline{S}}}_2^+(a) \\ \hat{\underline{\underline{S}}}_1^-(\zeta) \\ \hat{\underline{\underline{S}}}_2^-(\zeta) \end{bmatrix} + \begin{bmatrix} \mathfrak{s}_1(a,w,\zeta) \\ \mathfrak{s}_2(a,w,\zeta) \\ \mathfrak{S}_1(\zeta,w,a) \\ \mathfrak{S}_2(\zeta,w,a) \end{bmatrix} \tag{5.47}$$

The equations contained in 5.47 are the *invariant imbedding relations* for $[a, \zeta]$. These relations hold for any internal depth $w, a \le w \le \zeta$. As a mnemonic aid, note that depth w, which appears on the left side of the equation, in imbedded within (lies between) the boundary depths a and ζ. *Invariant imbedding relations show how to compute the radiances at any internal depth, given the radiances incident onto a slab from above and below.*

The $\mathfrak{T}_{ij}(a,w,\zeta)$ are known as the *complete transmittances for downwelling radiance* for the combined slab $[a,\zeta]$. (This report follows the notation of *Hydrologic Optics*, which used script letters for complete operators; *Light and Water* used bold face only because a script font was not available.) Note that these matrices transmit downwelling radiances from the surface at depth a to depth w, accounting for the infinite series of internal reflections within the combined slabs. $\mathfrak{R}_{ij}(a,w,\zeta)$ are the *complete reflectances for downwelling radiance* incident onto the sea surface. Likewise, $\mathfrak{T}_{ij}(\zeta,w,a)$ are the complete transmittances for upwelling radiance incident onto the bottom of the combined slab, and $\mathfrak{R}_{ij}(\zeta,w,a)$ are the corresponding complete reflectances for upwelling radiance. The $\mathfrak{s}_{1,2}(a,w,\zeta)$ are the *complete source-induced downwelling radiances*, and

$\mathfrak{S}_{1,2}(\zeta, w, a)$ are the *complete source-induced upwelling radiances*. The order of (a, w, ζ) or (ζ, w, a) shows the direction of the incident or internal-source radiance. The subscripts tell how cosine and sine amplitudes are transformed as they are transmitted or reflected. Thus $\mathfrak{T}_{11}(a, w, \zeta)$ transmits cosine amplitudes from a to w, $\mathfrak{T}_{12}(a, w, \zeta)$ transmits from a to w and turns sine amplitudes into cosine amplitudes in the process, $\mathfrak{T}_{21}(\zeta, w, a)$ transmits from ζ to w and converts cosine amplitudes to sine amplitudes, and so on.

Writing out the complete reflectances, transmittances, and source matrices in terms of the standard matrices gives (L&W 7.67-7.69)

$$\begin{aligned}
\mathfrak{T}_{11}(\zeta, w, a) &= [I - R_1 r_1 + R_2 r_2 (I - R_1 r_2)^{-1} R_2 r_1]^{-1} T_1 \\
&\quad - (I - R_1 r_1)^{-1} R_2 r_2 [I - R_1 r_2 + R_2 r_1 (I - R_1 r_1)^{-1} R_2 r_2]^{-1} T_2 && \text{(5.48a)} \\
\mathfrak{T}_{12}(\zeta, w, a) &= -[I - R_1 r_1 + R_2 r_2 (I - R_1 r_2)^{-1} R_2 r_1]^{-1} T_2 \\
&\quad - (I - R_1 r_1)^{-1} R_2 r_2 [I - R_1 r_2 + R_2 r_1 (I - R_1 r_1)^{-1} R_2 r_2]^{-1} T_1 && \text{(5.48b)} \\
\mathfrak{T}_{21}(\zeta, w, a) &= (I - R_1 r_2)^{-1} R_2 r_1 [I - R_1 r_1 + R_2 r_2 (I - R_1 r_2)^{-1} R_2 r_1]^{-1} T_1 \\
&\quad + [I - R_1 r_2 + R_2 r_1 (I - R_1 r_1)^{-1} R_2 r_2]^{-1} T_2 && \text{(5.48c)} \\
\mathfrak{T}_{22}(\zeta, w, a) &= -(I - R_1 r_2)^{-1} R_2 r_1 [I - R_1 r_1 + R_2 r_2 (I - R_1 r_2)^{-1} R_2 r_1]^{-1} T_2 \\
&\quad + [I - R_1 r_2 + R_2 r_1 (I - R_1 r_1)^{-1} R_2 r_2]^{-1} T_1 && \text{(5.48d)} \\
\mathfrak{R}_{11}(\zeta, w, a) &= r_1 \mathfrak{T}_{11}(\zeta, w, a) && \text{(5.48e)} \\
\mathfrak{R}_{12}(\zeta, w, a) &= r_1 \mathfrak{T}_{12}(\zeta, w, a) && \text{(5.48f)} \\
\mathfrak{R}_{21}(\zeta, w, a) &= r_2 \mathfrak{T}_{21}(\zeta, w, a) && \text{(5.48g)} \\
\mathfrak{R}_{22}(\zeta, w, a) &= r_2 \mathfrak{T}_{22}(\zeta, w, a) && \text{(5.48h)} \\
\mathfrak{S}_1(\zeta, w, a) &= (R_1 m_{11} - R_2 m_{21})\sigma_1 + (R_1 m_{12} - R_2 m_{22})\sigma_2 + m_{33}\Sigma_1 + m_{34}\Sigma_2 && \text{(5.48i)} \\
\mathfrak{S}_2(\zeta, w, a) &= (R_2 m_{11} + R_1 m_{21})\sigma_1 + (R_2 m_{12} + R_1 m_{22})\sigma_2 + m_{43}\Sigma_1 + m_{44}\Sigma_2 && \text{(5.48j)}
\end{aligned}$$

and (L&W 7.70-7.72):

$$\begin{aligned}
\mathfrak{T}_{11}(a, w, \zeta) &= [I - r_1 R_1 + r_1 R_2 (I - r_2 R_1)^{-1} r_2 R_2]^{-1} t_1 && \text{(5.49a)} \\
\mathfrak{T}_{12}(a, w, \zeta) &= -(I - r_1 R_1)^{-1} r_1 R_2 [I - r_2 R_2 + r_2 R_2 (I - r_1 R_1)^{-1} r_1 R_2]^{-1} t_2 && \text{(5.49b)} \\
\mathfrak{T}_{21}(a, w, \zeta) &= [(I - r_2 R_1)^{-1} r_2 R_2 [I - r_1 R_1 + r_1 R_2 (I - r_2 R_1)^{-1} r_2 R_2] t_1 && \text{(5.49c)} \\
\mathfrak{T}_{22}(a, w, \zeta) &= [I - r_2 R_1 + r_2 R_2 (I - r_1 R_1)^{-1} r_1 R_2]^{-1} t_2 && \text{(5.49d)} \\
\mathfrak{R}_{11}(a, w, \zeta) &= R_1 \mathfrak{T}_{11}(a, w, \zeta) - R_2 \mathfrak{T}_{21}(a, w, \zeta) && \text{(5.49e)} \\
\mathfrak{R}_{12}(a, w, \zeta) &= R_1 \mathfrak{T}_{12}(a, w, \zeta) - R_2 \mathfrak{T}_{22}(a, w, \zeta) && \text{(5.49f)} \\
\mathfrak{R}_{21}(a, w, \zeta) &= R_2 \mathfrak{T}_{11}(a, w, \zeta) + R_1 \mathfrak{T}_{21}(a, w, \zeta) && \text{(5.49g)} \\
\mathfrak{R}_{22}(a, w, \zeta) &= R_2 \mathfrak{T}_{12}(a, w, \zeta) + R_1 \mathfrak{T}_{22}(a, w, \zeta) && \text{(5.49h)} \\
\mathfrak{s}_1(a, w, \zeta) &= m_{11}\sigma_1 + m_{12}\sigma_2 + r_1 m_{33}\Sigma_1 + r_1 m_{34}\Sigma_2 && \text{(5.49i)} \\
\mathfrak{s}_2(a, w, \zeta) &= m_{21}\sigma_1 + m_{22}\sigma_2 + r_2 m_{43}\Sigma_1 + r_2 m_{44}\Sigma_2 && \text{(5.49j)}
\end{aligned}$$

The source terms are left in terms of the m_{ij} matrices simply because these equations are too lengthy to write out in full.

These equations are the *imbed rules* for the combined slabs $[a, w]$ and $[w, \zeta]$. *Imbed rules show how standard operators for two slabs are combined to generate complete operators for the composite*

slab.

A check on these equations can be obtained, as always, by considering the case of no polarization. Then R_2 and T_2 are zero within the water body (because the scalar phase functions has only cosine amplitudes), but r_2 and t_2 for the surface are not generally zero (because the surface has both sine and cosine amplitudes). For no polarization, $\mathfrak{T}_{11}(\zeta, w, a) = [I - R_1 r_1]^{-1} T_1$, which is (the transpose of) L&W Eq. (7.67). $\mathfrak{T}_{11}(a, w, \zeta)$ reduces to $[I - r_1 R_1]^{-1} t_1$, which corresponds to L&W (7.70). The "cross terms", $\mathfrak{T}_{12}(a, w, \zeta), \mathfrak{T}_{21}(a, w, \zeta), \mathfrak{R}_{12}(a, w, \zeta), \mathfrak{R}_{21}(a, w, \zeta)$ all reduce to zero and have no equivalent in the unpolarized case. However, $\mathfrak{T}_{22}(\zeta, w, a) = [I - R_1 r_2]^{-1} T_1$, which is almost, but not quite, the transpose of L&W (7.67); and $\mathfrak{T}_{22}(a, w, \zeta) = [I - r_2 R_1]^{-1} t_2$, which is almost equivalent to L&W (8.96). I am therefore not sure if the $\mathfrak{T}_{22}$ terms are correct; further checking is needed.

5.3.2 Interpretation of the Complete Operators

It is worthwhile to consider the physical interpretation of the complete operators defined in Eqs. (5.48) and (5.49). Consider, for example, $\mathfrak{T}_{11}(a, w, \zeta)$, which according to Eq. (5.47) converts $\underline{\hat{S}}_1^+(a)$ into $\underline{\hat{S}}_1^+(w)$.

$$\mathfrak{T}_{11}(a, w, \zeta) = \Big(\underbrace{I - r_1 R_1}_{A} + \underbrace{r_1 R_2 (I - r_2 R_1)^{-1} r_2 R_2}_{B} \Big)^{-1} t_1 \tag{5.50}$$

This inverse can be expanded using the formula

$$(A + B)^{-1} = A^{-1} - A^{-1} B (I + A^{-1} B)^{-1} A^{-1}$$

for the inverse of the sum of two matrices. Applying this expansion to $\mathfrak{T}_{11}(a, w, \zeta)$ with A and B as identified in Eq. (5.50) gives

$$\begin{aligned} \mathfrak{T}_{11}(a, w, \zeta) &= (I - r_1 R_1)^{-1} t_1 \\ &- (I - r_1 R_1)^{-1} r_1 R_2 (I - r_2 R_1)^{-1} r_2 R_2 \Big(I + (I - r_1 R_1)^{-1} r_1 R_2 (I - r_2 R_1)^{-1} r_2 R_2 \Big) (I - r_1 R_1)^{-1} t_1 \end{aligned} \tag{5.51}$$

Now consider the first term on the right hand side of this equation. Expanding the inverse of $(I - r_1 R_1)$ using the formula

$$(I - A)^{-1} = I + A + AA + AAA + \cdots$$

for the inverse of the identity minus a matrix, and reinserting the depth arguments, gives

$$\begin{aligned} \Big(I - r_1(a, w) R_1(w, \zeta) \Big)^{-1} t_1(a, w) &= t_1(a, w) + r_1(a, w) R_1(w, \zeta) t_1(a, w) \\ &+ r_1(a, w) R_1(w, \zeta) r_1(a, w) R_1(w, \zeta) t_1(a, w) + \cdots \end{aligned} \tag{5.52}$$

The upper panel of Fig. 5.1 shows the meaning of these terms when $\mathfrak{T}_{11}(a, w, \zeta)$ operates on the incident radiance $\underline{\hat{S}}_1^+(a)$. (This figure is similar to L&W Fig. 7.1, which includes the bottom boundary.) The first term of the expansion, $t_1(a, w)$ simply transmits $\underline{\hat{S}}_1^+(a)$ from above the sea

surface at depth a to just below the surface at depth w. The second term describes radiance that is transmitted through the surface by $t_1(a,w)$, then reflected upward by the water between w and depth ζ by the operation of $R_1(w,\zeta)$. This upward reflection creates upwelling radiance $\underline{\hat{S}}_1^-(w)$ at depth w. That radiance is then reflected back downward by the air-water surface via the operation of $r_1(w,a)$. The end result is another contribution to $\underline{\hat{S}}_1^+(w)$. The next term of the expansion represents another cycle of upward reflection by the water body followed by downward reflection by the sea surface.

Now consider the expansion of the second term on the right hand side of Eq. (5.51). Expanding the $(I - r_i R_j)^{-1}$ terms leads to the sequence of terms

$$\begin{aligned}\text{rhs second term} = & -r_1(w,a)R_2(w,\zeta)r_2(w,a)R_2(w,\zeta)t_1(a,w) \\ & - r_1(w,a)R_2(w,\zeta)r_2(w,a)R_1(w,\zeta)r_1(w,a)R_2(w,\zeta)r_2(w,a)R_1(w,\zeta)t_1(a,w) + \cdots\end{aligned}$$

The bottom panel of Fig. 5.1 shows the meaning of the first group of these terms. First $t_1(a,w)$ transmits $\underline{\hat{S}}_1^+(a)$ through the sea surface as before. Then reflection by $R_2(w,\zeta)$ converts the downwelling cosine amplitude $\underline{\hat{S}}_1^+(w)$ into an upwelling sine amplitude $\underline{\hat{S}}_2^-(w)$; recall that the standard operators for the water body R_2 and T_2 convert cosine amplitudes to sine amplitudes and vice versa. $\underline{\hat{S}}_2^-(w)$ is then reflected back downward by the water surface by $r_2(w,a)$ to create $\underline{\hat{S}}_2^+(w)$. That downwelling sine amplitude is then reflected back upward by the water body via $R_2(w,\zeta)$, which converts the sine back to a cosine amplitude $\underline{\hat{S}}_1^-(w)$, which is reflected back downward by $r_1(w,a)$. The end result is another contribution to $\underline{\hat{S}}_1^+(w)$. These terms do not occur is the SRTE because $R_2 = 0$, but in the VRTE they describe conversions between various components of the Stokes vectors by the elements of the phase matrix.

As another example, consider $\mathfrak{R}_{11}(\zeta,w,a)$, which according to Eq. (5.47) converts $\underline{\hat{S}}_1^-(\zeta)$ into $\underline{\hat{S}}_1^+(w)$. Performing a similar expansion of terms gives

$$\begin{aligned}\mathfrak{R}_{11}(\zeta,w,a) &= r_1[I - R_1 r_1 + R_2 r_2 (I - R_1 r_2)^{-1} R_2 r_1]^{-1} T_1 \\ &= r_1(w,a)T_1(\zeta,w) + r_1(w,a)R_1(w,\zeta)r_1(w,a)T_1(\zeta,w) + \cdots \\ &\quad - r_1(w,a)R_2(\zeta,w)r_2(w,a)T_2(\zeta,w) \\ &\quad - r_1(w,a)R_2(w,\zeta)r_2(w,a)R_1(w,\zeta)r_2(w,a)T_2(\zeta,w) + \cdots\end{aligned}$$

The first series of terms is equivalent to the corresponding terms in the SRTE: $T_1(\zeta,w)$ transmits upwelling radiance $\underline{\hat{S}}_1^-(\zeta)$ from ζ to the bottom of the sea surface, where it is reflected back down by $r_1(w,a)$ to create $\underline{\hat{S}}_1^+(w)$, and so on. Figure 5.2 shows the effect of the second group of terms involving $T_2(\zeta,w)$. Now $T_2(\zeta,w)$ transmits radiance from ζ to w, but converts cosine amplitudes to sine amplitudes. These are then reflected back down by $r_2(w,a)$. $R_1(\zeta,w)$ then reflects the downwelling $\underline{\hat{S}}_2^+(w)$ back upward, leaving it as a sine amplitude. Another reflection by $r_2(w,a)$ gives a downwelling sign amplitude, which is the reflected back upward by the water body and converted back to a cosine amplitude in the process. A final reflection by the sea surface gives another contribution to $\underline{\hat{S}}_1^+(w)$. This process does not occur in the SRTE because $T_2 = 0$.

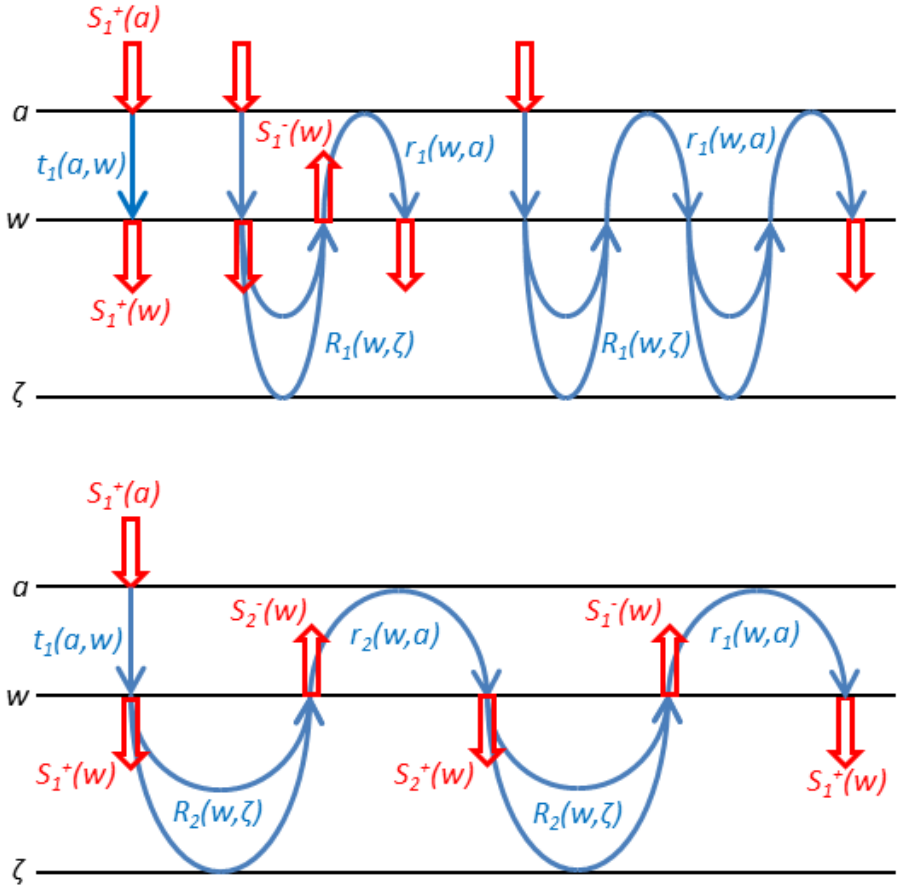

Figure 5.1: Graphical interpretation of the leading terms in the expansion of $\mathfrak{T}_{11}(a, w, \zeta)$.

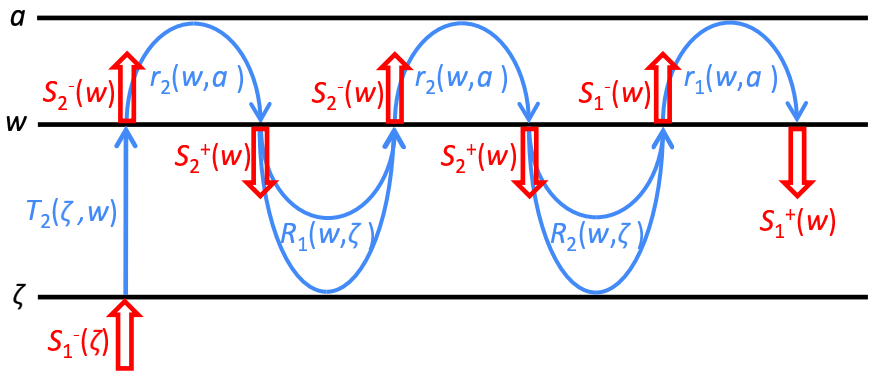

Figure 5.2: Graphical interpretation of one group of terms in the expansion of $\mathfrak{R}_{11}(\zeta, w, a)$.

The other complete operators has similar interpretations. *These infinite series of inter-reflections by the water body and the sea surface shows how invariant imbedding theory automatically accounts for all orders of multiple scattering within a water body and between the water and the surface (or bottom) boundary.* This result can be contrasted with the successive-order-scattering solution method. That method solves first for terms that represent unscattered radiance, then radiance scattered once, then radiance scattered twice, and so on. The solution process is then terminated after some number n of scatterings (often n is just 2 or 3). This gives an approximate solution of the RTE that represents the radiance due to a finite number of scatterings. That is, the successive-order-of-scattering solution approximates an infinite series of scatterings by adding up the first n terms. Invariant imbedding, on the other hand, solves directly for the analytical sum of the infinite series of scatterings.

5.3.3 Global Interaction Equations and Union Rules for the Surface Plus Water Body

We can now take the second step of the present derivation by using Eq. (5.47) to replace the internal radiances in Eqs. (5.18) with $p = 1, 2$, (5.22), and (5.24), which are (including for complete generality the possibility of internal sources in $[a, w]$)

$$\underline{\underline{\hat{S}}}_1^-(a) = \underline{\underline{\hat{r}}}_1(a,w)\underline{\underline{\hat{S}}}_1^+(a) + \underline{\underline{\hat{t}}}_1(w,a)\underline{\underline{\hat{S}}}_1^-(w) + \underline{\underline{\hat{\sigma}}}_1^{-t}(w,a)$$

$$\underline{\underline{\hat{S}}}_2^-(a) = \underline{\underline{\hat{r}}}_2(a,w)\underline{\underline{\hat{S}}}_2^+(a) + \underline{\underline{\hat{t}}}_2(w,a)\underline{\underline{\hat{S}}}_2^-(w) + \underline{\underline{\hat{\sigma}}}_2^{-t}(w,a)$$

$$\underline{\underline{\hat{S}}}_1^+(\zeta) = \underline{\underline{\hat{R}}}_1(\zeta,w)\underline{\underline{\hat{S}}}_1^-(\zeta) + \underline{\underline{\hat{T}}}_1(w,\zeta)\underline{\underline{\hat{S}}}_1^+(w) - \underline{\underline{\hat{R}}}_2(\zeta,w)\underline{\underline{\hat{S}}}_2^-(\zeta) - \underline{\underline{\hat{T}}}_2(w,\zeta)\underline{\underline{\hat{S}}}_2^+(w) + \underline{\underline{\hat{\Sigma}}}_1^{+t}(w,\zeta)$$

$$\underline{\underline{\hat{S}}}_2^+(\zeta) = \underline{\underline{\hat{R}}}_1(\zeta,w)\underline{\underline{\hat{S}}}_2^-(\zeta) + \underline{\underline{\hat{T}}}_1(w,\zeta)\underline{\underline{\hat{S}}}_2^+(w) + \underline{\underline{\hat{R}}}_2(\zeta,w)\underline{\underline{\hat{S}}}_1^-(\zeta) + \underline{\underline{\hat{T}}}_2(w,\zeta)\underline{\underline{\hat{S}}}_1^+(w) + \underline{\underline{\hat{\Sigma}}}_2^{+t}(w,\zeta)$$

Placing these equations in matrix form (again omitting the full matrix amplitude notation for brevity, so that $t_1(w,a)$ here represents $\underline{\underline{\hat{t}}}_1(w,a)$, etc.) gives

$$\begin{bmatrix} \underline{\underline{\hat{S}}}_1^-(a) \\ \underline{\underline{\hat{S}}}_2^-(a) \\ \underline{\underline{\hat{S}}}_1^+(\zeta) \\ \underline{\underline{\hat{S}}}_2^+(\zeta) \end{bmatrix} = \begin{bmatrix} 0 & 0 & t_1(w,a) & 0 \\ 0 & 0 & 0 & t_2(w,a) \\ T_1(w,\zeta) & -T_2(w,\zeta) & 0 & 0 \\ T_2(w,\zeta) & T_1(w,\zeta) & 0 & 0 \end{bmatrix} \begin{bmatrix} \underline{\underline{\hat{S}}}_1^+(w) \\ \underline{\underline{\hat{S}}}_2^+(w) \\ \underline{\underline{\hat{S}}}_1^-(w) \\ \underline{\underline{\hat{S}}}_2^-(w) \end{bmatrix}$$

$$+ \begin{bmatrix} r_1(a,w) & 0 & 0 & 0 \\ 0 & r_2(a,w) & 0 & 0 \\ 0 & 0 & R_1(\zeta,w) & -R_2(\zeta,w) \\ 0 & 0 & R_2(\zeta,w) & R_1(\zeta,w) \end{bmatrix} \begin{bmatrix} \underline{\underline{\hat{S}}}_1^+(a) \\ \underline{\underline{\hat{S}}}_2^+(a) \\ \underline{\underline{\hat{S}}}_1^-(\zeta) \\ \underline{\underline{\hat{S}}}_2^-(\zeta) \end{bmatrix} + \begin{bmatrix} \underline{\underline{\hat{\sigma}}}_1^{-t}(w,a) \\ \underline{\underline{\hat{\sigma}}}_2^{-t}(w,a) \\ \underline{\underline{\hat{\Sigma}}}_1^{+t}(w,\zeta) \\ \underline{\underline{\hat{\Sigma}}}_2^{+t}(w,\zeta) \end{bmatrix}$$

The invariant imbedding Eq. (5.47) is now used to replace the column vector of internal radiances at depth w in this equation. Carrying out the matrix multiplications gives (L&W 7.73, 7.74)

$$\begin{bmatrix} \hat{\underline{\underline{S}}}_1^-(a) \\ \hat{\underline{\underline{S}}}_2^-(a) \\ \hat{\underline{\underline{S}}}_1^+(\zeta) \\ \hat{\underline{\underline{S}}}_2^+(\zeta) \end{bmatrix} = \begin{bmatrix} t_1(w,a)\mathfrak{R}_{11}(a,w,\zeta) + r_1(a,w) & t_1(w,a)\mathfrak{R}_{12}(a,w,\zeta) \\ t_2(w,a)\mathfrak{R}_{21}(a,w,\zeta) & t_2(w,a)\mathfrak{R}_{22}(a,w,\zeta) + r_2(a,w) \\ T_1(w,\zeta)\mathfrak{T}_{11}(a,w,\zeta) - T_2(w,\zeta)\mathfrak{T}_{21}(a,w,\zeta) & T_1(w,\zeta)\mathfrak{T}_{12}(a,w,\zeta) - T_2(w,\zeta)\mathfrak{T}_{22}(a,w,\zeta) \\ T_2(w,\zeta)\mathfrak{T}_{11}(a,w,\zeta) + T_1(w,\zeta)\mathfrak{T}_{21}(a,w,\zeta) & T_2(w,\zeta)\mathfrak{T}_{12}(a,w,\zeta) + T_1(w,\zeta)\mathfrak{T}_{22}(a,w,\zeta) \end{bmatrix}$$

$$\begin{bmatrix} t_1(w,a)\mathfrak{T}_{11}(\zeta,w,a) \\ t_2(w,a)\mathfrak{T}_{21}(\zeta,w,a) \\ T_1(w,\zeta)\mathfrak{R}_{11}(\zeta,w,a) - T_2(w,\zeta)\mathfrak{R}_{21}(\zeta,w,a) + R_1(\zeta,w) \\ T_2(w,\zeta)\mathfrak{R}_{11}(\zeta,w,a) + T_1(w,\zeta)\mathfrak{R}_{21}(\zeta,w,a) + R_2(\zeta,w) \end{bmatrix}$$

$$\begin{bmatrix} t_1(w,a)\mathfrak{T}_{12}(\zeta,w,a) \\ t_2(w,a)\mathfrak{T}_{22}(\zeta,w,a) \\ T_1(w,\zeta)\mathfrak{R}_{12}(\zeta,w,a) - T_2(w,\zeta)\mathfrak{R}_{22}(\zeta,w,a) - R_2(\zeta,w) \\ T_2(w,\zeta)\mathfrak{R}_{12}(a,w,\zeta) + T_1(w,\zeta)\mathfrak{R}_{22}(a,w,\zeta) + R_1(\zeta,w) \end{bmatrix} \begin{bmatrix} \hat{\underline{\underline{S}}}_1^+(a) \\ \hat{\underline{\underline{S}}}_2^+(a) \\ \hat{\underline{\underline{S}}}_1^-(\zeta) \\ \hat{\underline{\underline{S}}}_2^-(\zeta) \end{bmatrix}$$

$$+ \begin{bmatrix} t_1(w,a)\mathfrak{S}_1(\zeta,w,a) + \hat{\underline{\underline{\sigma}}}_1^{-t}(w,a) \\ t_2(w,a)\mathfrak{S}_2(\zeta,w,a) + \hat{\underline{\underline{\sigma}}}_2^{-t}(w,a) \\ T_1(w,\zeta)\mathfrak{s}_1(a,w,\zeta) - T_2(w,\zeta)\mathfrak{s}_2(a,w,\zeta) + \hat{\underline{\underline{\Sigma}}}_1^{+t}(w,\zeta) \\ T_2(w,\zeta)\mathfrak{s}_1(a,w,\zeta) + T_1(w,\zeta)\mathfrak{s}_2(a,w,\zeta) + \hat{\underline{\underline{\Sigma}}}_2^{+t}(w,\zeta) \end{bmatrix} \tag{5.53}$$

These equations show how the composite slab $[a,\zeta]$ responds to incident radiances; they are therefore the *global interaction equations for the composite slab.* Global interaction equations are written in terms of standard operators, e.g. as in Eqs. (5.21)-(5.24). We therefore rewrite the last equation as

$$\begin{bmatrix} \hat{\underline{\underline{S}}}_1^-(a) \\ \hat{\underline{\underline{S}}}_2^-(a) \\ \hat{\underline{\underline{S}}}_1^+(\zeta) \\ \hat{\underline{\underline{S}}}_2^+(\zeta) \end{bmatrix} \equiv \begin{bmatrix} R_{11}(a,\zeta) & R_{12}(a,\zeta) & T_{11}(\zeta,a) & T_{12}(\zeta,a) \\ R_{21}(a,\zeta) & R_{22}(a,\zeta) & T_{21}(\zeta,a) & T_{22}(\zeta,a) \\ T_{11}(a,\zeta) & T_{12}(a,\zeta) & R_{11}(\zeta,a) & R_{12}(\zeta,a) \\ T_{21}(a,\zeta) & T_{22}(a,\zeta) & R_{21}(\zeta,a) & R_{22}(\zeta,a) \end{bmatrix} \begin{bmatrix} \hat{\underline{\underline{S}}}_1^+(a) \\ \hat{\underline{\underline{S}}}_2^+(a) \\ \hat{\underline{\underline{S}}}_1^-(\zeta) \\ \hat{\underline{\underline{S}}}_2^-(\zeta) \end{bmatrix} + \begin{bmatrix} \hat{\underline{\underline{\sigma}}}_1^{-t}(\zeta,a) \\ \hat{\underline{\underline{\sigma}}}_2^{-t}(\zeta,a) \\ \hat{\underline{\underline{\Sigma}}}_1^{+t}(a,\zeta) \\ \hat{\underline{\underline{\Sigma}}}_2^{+t}(a,\zeta) \end{bmatrix} \tag{5.54}$$

The equations defining the standard operators for the combined slab are (L&W 7.75-7.77),

$$R_{11}(a,\zeta) = t_1(w,a)\mathfrak{R}_{11}(a,w,\zeta) + r_1(a,w) \tag{5.55}$$
$$R_{12}(a,\zeta) = t_1(w,a)\mathfrak{R}_{12}(a,w,\zeta) \tag{5.56}$$
$$R_{21}(a,\zeta) = t_2(w,a)\mathfrak{R}_{21}(a,w,\zeta) \tag{5.57}$$
$$R_{22}(a,\zeta) = t_2(w,a)\mathfrak{R}_{22}(a,w,\zeta) + r_2(a,w) \tag{5.58}$$
$$T_{11}(\zeta,a) = t_1(w,a)\mathfrak{T}_{11}(\zeta,w,a) \tag{5.59}$$
$$T_{12}(\zeta,a) = t_1(w,a)\mathfrak{T}_{12}(\zeta,w,a) \tag{5.60}$$
$$T_{21}(\zeta,a) = t_2(w,a)\mathfrak{T}_{21}(\zeta,w,a) \tag{5.61}$$
$$T_{22}(\zeta,a) = t_2(w,a)\mathfrak{T}_{22}(\zeta,w,a) \tag{5.62}$$
$$\hat{\underline{\underline{\sigma}}}_1^{-t}(\zeta,a) = t_1(w,a)\mathfrak{S}_1(\zeta,w,a) + \hat{\underline{\underline{\sigma}}}_1^{-t}(w,a) \tag{5.63}$$
$$\hat{\underline{\underline{\sigma}}}_2^{-t}(\zeta,a) = t_2(w,a)\mathfrak{S}_2(\zeta,w,a) + \hat{\underline{\underline{\sigma}}}_2^{-t}(w,a) \tag{5.64}$$

and (L&W 7.78-7.80)

$$R_{11}(\zeta,a) = T_1(w,\zeta)\mathfrak{R}_{11}(\zeta,w,a) - T_2(w,\zeta)\mathfrak{R}_{21}(\zeta,w,a) + R_1(\zeta,w) \tag{5.65}$$
$$R_{12}(\zeta,a) = T_1(w,\zeta)\mathfrak{R}_{12}(\zeta,w,a) - T_2(w,\zeta)\mathfrak{R}_{22}(\zeta,w,a) - R_2(\zeta,w) \tag{5.66}$$
$$R_{21}(\zeta,a) = T_2(w,\zeta)\mathfrak{R}_{11}(\zeta,w,a) + T_1(w,\zeta)\mathfrak{R}_{21}(\zeta,w,a) + R_2(\zeta,w) \tag{5.67}$$
$$R_{22}(\zeta,a) = T_2(w,\zeta)\mathfrak{R}_{12}(\zeta,w,a) + T_1(w,\zeta)\mathfrak{R}_{22}(\zeta,w,a) + R_1(\zeta,w) \tag{5.68}$$
$$T_{11}(a,\zeta) = T_1(w,\zeta)\mathfrak{T}_{11}(a,w,\zeta) - T_2(w,\zeta)\mathfrak{T}_{21}(a,w,\zeta) \tag{5.69}$$
$$T_{12}(a,\zeta) = T_1(w,\zeta)\mathfrak{T}_{12}(a,w,\zeta) - T_2(w,\zeta)\mathfrak{T}_{22}(a,w,\zeta) \tag{5.70}$$
$$T_{21}(a,\zeta) = T_2(w,\zeta)\mathfrak{T}_{11}(a,w,\zeta) + T_1(w,\zeta)\mathfrak{T}_{21}(a,w,\zeta) \tag{5.71}$$
$$T_{22}(a,\zeta) = T_2(w,\zeta)\mathfrak{T}_{12}(a,w,\zeta) + T_1(w,\zeta)\mathfrak{T}_{22}(a,w,\zeta) \tag{5.72}$$
$$\hat{\underline{\underline{\Sigma}}}_1^{+t}(a,\zeta) = T_1(w,\zeta)\mathfrak{s}_1(a,w,\zeta) - T_2(w,\zeta)\mathfrak{s}_2(a,w,\zeta) + \hat{\underline{\underline{\Sigma}}}_1^{+t}(w,\zeta) \tag{5.73}$$
$$\hat{\underline{\underline{\Sigma}}}_2^{+t}(a,\zeta) = T_2(w,\zeta)\mathfrak{s}_1(a,w,\zeta) + T_1(w,\zeta)\mathfrak{s}_2(a,w,\zeta) + \hat{\underline{\underline{\Sigma}}}_2^{+t}(w,\zeta) \tag{5.74}$$

Recalling that the imbed rules (5.48) and (5.48) define the complete operators in terms of standard operators, these equations show how the standard operators for two slabs are combined to obtain the standard operators for a composite slab. We can imagine holding a and w fixed and "constructing" the water body by letting ζ increase downward starting at $\zeta = w$. These relations are therefore known as the *downward union rules* for the composite slab.

As always, we can check that the VRTE equations reduce to their SRTE equivalents. Thus $T_{11}(\zeta,a)$ reduces to (the transpose of) L&W (7.75), $R_{11}(a,\zeta)$ is the equivalent of L&W (7.76), $R_{12}(a,\zeta) = R_{21}(a,\zeta) = 0$ and have no scalar equivalents, and $\hat{\underline{\underline{\sigma}}}_1^{-t}(\zeta,a)$ reduces to L&W(7.77). Likewise, $R_{11}(\zeta,a)$ reduces to L&W(7.79), $T_{11}(a,\zeta)$ is equivalent to L&W(7.78), and so on[1].

[1] As was the case for the $\mathfrak{T}_{22}$ complete transmittances, the reductions of the T_{22} standard matrices to the unpolarized case are not obvious; need to check further.

A review the preceding developments is warranted.

- The transfer functions (standard operators) for the air-water surface, $r_{1,2}(a,w)$, $r_{1,2}(w,a)$, $t_{1,2}(a,w)$, and $t_{1,2}(w,a)$, are known from the surface boundary conditions.
- The standard operators for the bare slab $[w,\zeta]$, $R_1(w,\zeta)$, etc. are known from the downward sweep integrations of the ODEs (4.48-4.54) beginning with initial conditions (4.55-4.58)
- These two sets of standard operators were combined via the imbed rules (5.47) and (5.48-5.49). Those equations give the radiances at depth w, $S^{\pm}_{1,2}(w)$, in terms of the incident radiances at depths a and ζ. However, the radiances at ζ are not yet known.
- The global interaction equations (5.53, 5.54) and union rules (5.64, 5.74) give the standard operators $R_{11}(a,\zeta)$ etc. for the combined slabs $[a,w]\cup[w,\zeta]$.
- In general the above process would be repeated to combine the bare slab $[\zeta,m]$ with the lower boundary $[m,b]$. However, the lower boundary was assumed to be an azimuthally isotropic reflecting layer, in which case the Fourier modes decoupled ℓ value by ℓ value. This allowed the standard operators $R_{1,2}(\zeta,b)$ etc. for the slab $[\zeta,m]\cup[m,b]$ to be obtained by the upward sweep integrations of the ODEs (4.59-4.70) with initial conditions (5.36-5.39) for each ℓ value. Thus the standard operators $R_1(\zeta,b)$ etc. are known from the upward sweep ODE solutions and the known bottom reflectance.

5.4 Computing the Radiances at Depth w

The next step is to combine slabs $[a,\zeta]$, and $[\zeta,b]$, both of whose properties are now known and incorporate the surface and bottom boundary conditions. It will be convenient first to let $\zeta = w$ so as to obtain the upward and downward radiances at depth w. We will then use those values and the standard operators for slabs $[w,\zeta]$ and $[\zeta,b]$ to obtain the radiances at any depth ζ within the water body. It will then be possible to obtain the water-leaving radiances. The final step of the solution is then to compute the physical-space Stokes vectors from the Fourier amplitudes.

The invariant imbedding equation for the depth triplet (a,w,b) can be written with the same form as Eq. (5.47):

$$\begin{bmatrix} \hat{\underline{\underline{S}}}_1^+(w) \\ \hat{\underline{\underline{S}}}_2^+(w) \\ \hat{\underline{\underline{S}}}_1^-(w) \\ \hat{\underline{\underline{S}}}_2^-(w) \end{bmatrix} \equiv \begin{bmatrix} \mathfrak{T}_{11}(a,w,b) & \mathfrak{T}_{12}(a,w,b) & \mathfrak{R}_{11}(b,w,a) & \mathfrak{R}_{12}(b,w,a) \\ \mathfrak{T}_{21}(a,w,b) & \mathfrak{T}_{22}(a,w,b) & \mathfrak{R}_{21}(b,w,a) & \mathfrak{R}_{22}(b,w,a) \\ \mathfrak{R}_{11}(a,w,b) & \mathfrak{R}_{12}(a,w,b) & \mathfrak{T}_{11}(b,w,a) & \mathfrak{T}_{12}(b,w,a) \\ \mathfrak{R}_{21}(a,w,b) & \mathfrak{R}_{22}(a,w,b) & \mathfrak{T}_{21}(b,w,a) & \mathfrak{T}_{22}(b,w,a) \end{bmatrix} \begin{bmatrix} \hat{\underline{\underline{S}}}_1^+(a) \\ \hat{\underline{\underline{S}}}_2^+(a) \\ \hat{\underline{\underline{S}}}_1^-(b) \\ \hat{\underline{\underline{S}}}_2^-(b) \end{bmatrix} + \begin{bmatrix} \mathfrak{s}_1(a,w,b) \\ \mathfrak{s}_2(a,w,b) \\ \mathfrak{S}_1(b,w,a) \\ \mathfrak{S}_2(b,w,a) \end{bmatrix} \tag{5.75}$$

Now, however, the upwelling radiance at depth b is zero because no light enters the water from below the bottom boundary. Thus $\hat{\underline{\underline{S}}}^-_{1,2}(b) = 0$, and these equations reduce to (L&W 8.95 and 8.99

with $p = 1$ and 2)

$$\underline{\underline{\hat{S}}}_1^+(w) = \mathfrak{T}_{11}(a,w,b)\underline{\underline{\hat{S}}}_1^+(a) + \mathfrak{T}_{12}(a,w,b)\underline{\underline{\hat{S}}}_2^+(a) + \mathfrak{s}_1(a,w,b) \tag{5.76}$$

$$\underline{\underline{\hat{S}}}_2^+(w) = \mathfrak{T}_{21}(a,w,b)\underline{\underline{\hat{S}}}_1^+(a) + \mathfrak{T}_{22}(a,w,b)\underline{\underline{\hat{S}}}_2^+(a) + \mathfrak{s}_2(a,w,b) \tag{5.77}$$

$$\underline{\underline{\hat{S}}}_1^-(w) = \mathfrak{R}_{11}(a,w,b)\underline{\underline{\hat{S}}}_1^+(a) + \mathfrak{R}_{12}(a,w,b)\underline{\underline{\hat{S}}}_2^+(a) + \mathfrak{S}_1(b,w,a) \tag{5.78}$$

$$\underline{\underline{\hat{S}}}_2^-(w) = \mathfrak{R}_{21}(a,w,b)\underline{\underline{\hat{S}}}_1^+(a) + \mathfrak{R}_{22}(a,w,b)\underline{\underline{\hat{S}}}_2^+(a) + \mathfrak{S}_2(b,w,a) \tag{5.79}$$

The complete operators $\mathfrak{T}_{11}(a,w,b)$ etc. are given by equations of the same form as Eqs. (5.44-5.46). Thus in analogy to Eq. (5.49a) we have

$$\begin{aligned}\mathfrak{T}_{11}(a,w,b) = [&I - r_1(w,a)R_1(w,b) \\ &+ r_1(w,a)R_2(w,b)(I - r_2(w,a)R_1(w,b))^{-1}r_2(w,a)R_2(w,b)]^{-1}t_1(a,w)\end{aligned} \tag{5.80}$$

and so on.

For the case of no polarization, $R_2(w,b) = 0$ and Eq. (5.76) reduces to (L&W 8.95-8.97 for $p = 1$)

$$\begin{aligned}\underline{\underline{\hat{S}}}_1^+(w) = &[I - r_1(w,a)R_1(w,b)]^{-1}t_1(a,w)\underline{\underline{\hat{S}}}_1^+(a) \\ &+ [I - r_1(w,a)R_1(w,b)]^{-1}\sigma(a,w) + r_1(w,a)[I - R_1(w,b)r_1(w,a)]^{-1}\Sigma(b,w)\end{aligned}$$

Moreover, in this equation $\sigma(a,w) = 0$ because we are assuming that the air-water surface itself is source free.

Equations (5.76-5.79) give the radiances at depth w, $\underline{\underline{\hat{S}}}_{1,2}^{\pm}(w)$, in terms of the known incident radiances onto the sea surface, $\underline{\underline{\hat{S}}}_{1,2}^{+}(a)$, and the internal sources in the water body, $\mathfrak{S}_{1,2}(b,w,a)$.

5.5 Computing the Radiances at Depth ζ

Now that the radiances at depth w are known, they can be used along with the known standard operators for slab $[w,\zeta]$ and for slab $[\zeta,b]$ to write down global interaction equations for the combined slab $[w,\zeta] \cup [\zeta,b]$. These equations have the same form as the global interaction equations seen before in Eqs. (4.30, 4.32, 4.33, and 4.35), with the exception that depth m is replaced by b.

The result is

$$\hat{\underline{\underline{S}}}_1^+(\zeta|\ell) = \hat{\underline{\underline{R}}}_1(\zeta,w|\ell)\hat{\underline{\underline{S}}}_1^-(\zeta|\ell) + \hat{\underline{\underline{T}}}_1(w,\zeta|\ell)\hat{\underline{\underline{S}}}_1^+(w|\ell)$$
$$- \hat{\underline{\underline{R}}}_2(\zeta,w|\ell)\hat{\underline{\underline{S}}}_2^-(\zeta|\ell) - \hat{\underline{\underline{T}}}_2(w,\zeta|\ell)\hat{\underline{\underline{S}}}_2^+(w|\ell) + \hat{\underline{\underline{\Sigma}}}_1^{+t}(w,\zeta|\ell) \tag{5.81}$$

$$\hat{\underline{\underline{S}}}_2^+(\zeta|\ell) = \hat{\underline{\underline{R}}}_1(\zeta,w|\ell)\hat{\underline{\underline{S}}}_2^-(\zeta|\ell) + \hat{\underline{\underline{T}}}_1(w,\zeta|\ell)\hat{\underline{\underline{S}}}_2^+(w|\ell)$$
$$+ \hat{\underline{\underline{R}}}_2(\zeta,w|\ell)\hat{\underline{\underline{S}}}_1^-(\zeta|\ell) + \hat{\underline{\underline{T}}}_2(w,\zeta|\ell)\hat{\underline{\underline{S}}}_1^+(w|\ell) + \hat{\underline{\underline{\Sigma}}}_2^{+t}(w,\zeta|\ell) \tag{5.82}$$

$$\hat{\underline{\underline{S}}}_1^-(\zeta|\ell) = \hat{\underline{\underline{R}}}_1(\zeta,b|\ell)\hat{\underline{\underline{S}}}_1^+(\zeta|\ell) + \hat{\underline{\underline{T}}}_1(b,\zeta|\ell)\hat{\underline{\underline{S}}}_1^-(b|\ell)$$
$$- \hat{\underline{\underline{R}}}_2(\zeta,b|\ell)\hat{\underline{\underline{S}}}_2^+(\zeta|\ell) - \hat{\underline{\underline{T}}}_2(b,\zeta|\ell)\hat{\underline{\underline{S}}}_2^-(b|\ell) + \hat{\underline{\underline{\Sigma}}}_1^{-t}(b,\zeta|\ell) \tag{5.83}$$

$$\hat{\underline{\underline{S}}}_2^-(\zeta|\ell) = \hat{\underline{\underline{R}}}_1(\zeta,m|\ell)\hat{\underline{\underline{S}}}_2^+(\zeta|\ell) + \hat{\underline{\underline{T}}}_1(m,\zeta|\ell)\hat{\underline{\underline{S}}}_2^-(m|\ell)$$
$$+ \hat{\underline{\underline{R}}}_2(\zeta,b|\ell)\hat{\underline{\underline{S}}}_1^+(\zeta|\ell) + \hat{\underline{\underline{T}}}_2(b,\zeta|\ell)\hat{\underline{\underline{S}}}_1^-(b|\ell) + \hat{\underline{\underline{\Sigma}}}_2^{-t}(b,\zeta|\ell) \tag{5.84}$$

These four equations can be solved for each ℓ value for the internal radiances at any depth ζ in terms of the known incident radiances at depths w and b. For this development, we drop the ℓ argument and simplify the notation as before, so that $R_1(\zeta,w) = \hat{\underline{\underline{R}}}_1(\zeta,w|\ell)$, etc. Placing these equations in matrix form then gives the solution, which is similar in form to Eq. (5.41):

$$\begin{bmatrix} \hat{\underline{\underline{S}}}_1^+(\zeta) \\ \hat{\underline{\underline{S}}}_2^+(\zeta) \\ \hat{\underline{\underline{S}}}_1^-(\zeta) \\ \hat{\underline{\underline{S}}}_2^-(\zeta) \end{bmatrix} = \begin{bmatrix} I & 0 & -R_1(\zeta,w) & R_2(\zeta,w) \\ 0 & I & -R_2(\zeta,w) & -R_1(\zeta,w) \\ -R_1(\zeta,b) & R_2(\zeta,b) & I & 0 \\ -R_2(\zeta,b) & -R_1(\zeta,b) & 0 & I \end{bmatrix}^{-1} \times$$
$$\left\{ \begin{bmatrix} T_1(w,\zeta) & -T_2(w,\zeta) & 0 & 0 \\ T_2(w,\zeta) & T_1(w,\zeta) & 0 & 0 \\ 0 & 0 & T_1(b,\zeta) & -T_2(b,\zeta) \\ 0 & 0 & T_2(b,\zeta) & T_1(b,\zeta) \end{bmatrix} \begin{bmatrix} \hat{\underline{\underline{S}}}_1^+(w) \\ \hat{\underline{\underline{S}}}_2^+(w) \\ \hat{\underline{\underline{S}}}_1^-(b) \\ \hat{\underline{\underline{S}}}_2^-(b) \end{bmatrix} + \begin{bmatrix} \Sigma_1^{+t}(w,\zeta) \\ \Sigma_2^{+t}(w,\zeta) \\ \Sigma_1^{-t}(b,\zeta) \\ \Sigma_2^{-t}(b,\zeta) \end{bmatrix} \right\} \tag{5.85}$$

Repeated application of the block matrix inversion formulas (5.42) and (5.43) can again be used to obtain the inverse of the 4×4 block matrix. Letting B and C be the upper right and lower left 2×2 blocks respectively, we now have

$$(I - BC)^{-1} = \begin{bmatrix} I - R_1(\zeta,w)R_1(\zeta,b) + R_2(\zeta,w)R_2(\zeta,b) & R_1(\zeta,w)R_2(\zeta,b) + R_2(\zeta,w)R_1(\zeta,b) \\ -R_2(\zeta,w)R_1(\zeta,b) - R_1(\zeta,w)R_2(\zeta,b) & I + R_2(\zeta,w)R_2(\zeta,b) - R_1(\zeta,w)R_1(\zeta,b) \end{bmatrix}^{-1}$$
$$\equiv \begin{bmatrix} \alpha & \beta \\ \gamma & \delta \end{bmatrix}^{-1} = \begin{bmatrix} (\alpha - \beta\delta^{-1}\gamma)^{-1} & -\alpha^{-1}\beta(\delta - \gamma\alpha^{-1}\beta)^{-1} \\ \delta^{-1}\gamma(\alpha - \beta\delta^{-1}\gamma)^{-1} & (\delta - \gamma\alpha^{-1}\beta)^{-1} \end{bmatrix} \equiv \begin{bmatrix} n_{11} & n_{12} \\ n_{21} & n_{22} \end{bmatrix} \tag{5.86}$$

$$-C(I - BC)^{-1} = \begin{bmatrix} -R_1(\zeta,b)n_{11} + R_2(\zeta,b)n_{21} & -R_1(\zeta,b)n_{12} + R_2(\zeta,b)n_{22} \\ -R_2(\zeta,b)n_{11} - R_2(\zeta,b)n_{21} & -R_2(\zeta,b)n_{12} + R_1(\zeta,b)n_{22} \end{bmatrix} \tag{5.87}$$

$$(I-CB)^{-1}=\begin{bmatrix} I-R_1(\zeta,b)R_1(\zeta,w)+R_2(\zeta,b)R_2(\zeta,w) & R_1(\zeta,b)R_2(\zeta,w)+R_2(\zeta,b)R_1(\zeta,2) \\ -R_2(\zeta,b)R_1(\zeta,w)-R_1(\zeta,b)R_2(\zeta,w) & I+R_2(\zeta,b)R_2(\zeta,w)-R_1(\zeta,b)R_1(\zeta,w) \end{bmatrix}^{-1}$$
$$\equiv\begin{bmatrix} a & b \\ c & d \end{bmatrix}^{-1}=\begin{bmatrix} (a-bd^{-1}c)^{-1} & -a^{-1}b(d-ca^{-1}b)^{-1} \\ d^{-1}c(a-bd^{-1}c)^{-1} & (d-ca^{-1}b)^{-1} \end{bmatrix}\equiv\begin{bmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \end{bmatrix} \tag{5.88}$$

$$-B(I-CB)^{-1}=\begin{bmatrix} -R_1(\zeta,w)p_{11}+R_2(\zeta,w)p_{21} & -R_1(\zeta,w)p_{12}+R_2(\zeta,w)p_{22} \\ -R_2(\zeta,w)p_{11}-R_2(\zeta,w)p_{21} & -R_2(\zeta,w)p_{12}+R_1(\zeta,w)p_{22} \end{bmatrix} \tag{5.89}$$

Equation (5.43) now gives the inverse in Eq. (5.85) as

$$\begin{bmatrix} I & 0 & -R_1(\zeta,w) & R_2(\zeta,w) \\ 0 & I & -R_2(\zeta,w) & -R_1(\zeta,w) \\ -R_1(\zeta,b) & R_2(\zeta,b) & I & 0 \\ -R_2(\zeta,b) & -R_1(\zeta,b) & 0 & I \end{bmatrix}^{-1} =$$
$$\begin{bmatrix} & n_{11} & & n_{12} \\ & n_{21} & & n_{22} \\ -R_1(\zeta,b)n_{11}+R_2(\zeta,b)n_{21} & & -R_1(\zeta,b)n_{12}+R_2(\zeta,b)n_{22} & \\ -R_2(\zeta,b)n_{11}-R_2(\zeta,b)n_{21} & & -R_2(\zeta,b)n_{12}-R_1(\zeta,b)n_{22} & \end{bmatrix}$$

$$\begin{bmatrix} -R_1(\zeta,w)p_{11}+R_2(\zeta,w)p_{21} & -R_1(\zeta,w)p_{12}+R_2(\zeta,w)p_{22} \\ -R_2(\zeta,w)p_{11}-R_2(\zeta,w)p_{21} & -R_2(\zeta,w)p_{12}+R_1(\zeta,w)p_{22} \\ p_{11} & p_{12} \\ p_{21} & p_{22} \end{bmatrix} \tag{5.90}$$

Equation (5.85) now becomes

$$\begin{bmatrix} \hat{\underline{\underline{S}}}_1^+(\zeta) \\ \hat{\underline{\underline{S}}}_2^+(\zeta) \\ \hat{\underline{\underline{S}}}_1^-(\zeta) \\ \hat{\underline{\underline{S}}}_2^-(\zeta) \end{bmatrix} = \begin{bmatrix} n_{11}T_1(w,\zeta) + n_{12}T_2(w,\zeta) \\ n_{12}T_1(w,\zeta) + n_{22}T_2(w,\zeta) \\ [-R_1(\zeta,b)n_{11} + R_2(\zeta,b)n_{21}]T_1(w,\zeta) + [-R_1(\zeta,b)n_{12} + R_2(\zeta,b)n_{22}]T_2(w,\zeta) \\ [-R_2(\zeta,b)n_{11} - R_1(\zeta,b)n_{21}]T_1(w,\zeta) + [-R_2(\zeta,b)n_{12} - R_2(\zeta,b)n_{22}]T_2(w,\zeta) \end{bmatrix}$$

$$\begin{bmatrix} n_{11}T_2(w,\zeta) + n_{12}T_1(w,\zeta) \\ -n_{12}T_2(w,\zeta) + n_{22}T_1(w,\zeta) \\ [-R_1(\zeta,b)n_{11} + R_2(\zeta,b)n_{21}][-T_2(w,\zeta)] + [-R_1(\zeta,b)p_{12} + R_2(\zeta,b)p_{22}]T_2(w,\zeta) \\ [-R_2(\zeta,b)n_{11} - R_1(\zeta,b)n_{21}][-T_2(w,\zeta)] + [-R_1(\zeta,b)p_{12} + R_2(\zeta,b)p_{22}]T_2(w,\zeta) \end{bmatrix}$$

$$\begin{bmatrix} [-R_1(\zeta,w)p_{11} + R_2(\zeta,w)p_{21}]T_1(b,\zeta) + [-R_1(\zeta,w)p_{12} + R_2(\zeta,w)p_{22}]T_2(b,\zeta) \\ [-R_2(\zeta,w)p_{11} - R_1(\zeta,w)p_{21}]T_1(b,\zeta) + ([-R_2(\zeta,w)p_{12} - R_2(\zeta,w)p_{22}]T_2(b,\zeta) \\ p_{11}T_1(b,\zeta) + p_{12}T_2(b,\zeta) \\ p_{21}T_1(b,\zeta) + p_{22}T_2(b,\zeta) \end{bmatrix}$$

$$\begin{bmatrix} [-R_1(\zeta,w)p_{11} + R_2(\zeta,w)p_{21}][-T_2(b,\zeta)] + [-R_1(\zeta,w)p_{12} + R_2(\zeta,w)p_{22}]T_1(b,\zeta) \\ [-R_2(\zeta,w)p_{11} - R_1(\zeta,w)p_{21}][-T_2(b,\zeta)] + [-R_2(\zeta,w)p_{12} - R_2(\zeta,w)p_{22}]T_1(b,\zeta) \\ -p_{11}T_2(b,\zeta) + p_{12}T_1(b,\zeta) \\ -p_{21}T_2(b,\zeta) + p_{22}T_1(b,\zeta) \end{bmatrix} \begin{bmatrix} \hat{\underline{\underline{S}}}_1^+(w) \\ \hat{\underline{\underline{S}}}_2^+(w) \\ \hat{\underline{\underline{S}}}_1^-(b) \\ \hat{\underline{\underline{S}}}_2^-(b) \end{bmatrix}$$

$$+ \begin{bmatrix} n_{11}\Sigma_1^{+t}(w,\zeta) + n_{12}\Sigma_2^{+t}(w,\zeta) \\ n_{12}\Sigma_1^{+t}(w,\zeta) + n_{22}\Sigma_2^{+t}(w,\zeta) \\ [-R_1(\zeta,b)n_{11} + R_2(\zeta,b)n_{21}]\Sigma_1^{+t}(w,\zeta) + [-R_2(\zeta,b)n_{11} - R_1(\zeta,b)n_{21}]\Sigma_1^{+t}(w,\zeta) \\ [-R_2(\zeta,b)n_{11} - R_1(\zeta,b)n_{21}]\Sigma_1^{+t}(w,\zeta) + [-R_2(\zeta,b)n_{12} - R_1(\zeta,b)n_{21}]\Sigma_1^{+t}(w,\zeta) \end{bmatrix}$$

$$\begin{bmatrix} [-R_1(\zeta,w)p_{11} + R_2(\zeta,w)p_{21}]\Sigma_1^{-t}(b,\zeta) + [-R_1(\zeta,w)p_{12} + R_2(\zeta,w)p_{22}]\Sigma_2^{-t}(b,\zeta) \\ [-R_2(\zeta,w)p_{11} - R_1(\zeta,w)p_{21}]\Sigma_1^{-t}(b,\zeta) + [-R_2(\zeta,w)p_{12} - R_1(\zeta,w)p_{22}]\Sigma_2^{-t}(b,\zeta) \\ p_{11}\Sigma_1^{-t}(b,\zeta) + p_{12}\Sigma_2^{-t}(b,\zeta) \\ p_{21}\Sigma_1^{-t}(b,\zeta) + p_{22}\Sigma_2^{-t}(b,\zeta) \end{bmatrix} \tag{5.91}$$

Writing this in terms of complete operators gives

$$\begin{bmatrix} \hat{\underline{\underline{S}}}_1^+(\zeta) \\ \hat{\underline{\underline{S}}}_2^+(\zeta) \\ \hat{\underline{\underline{S}}}_1^-(\zeta) \\ \hat{\underline{\underline{S}}}_2^-(\zeta) \end{bmatrix} \equiv \begin{bmatrix} \mathfrak{T}_{11}(w,\zeta,b) & \mathfrak{T}_{12}(w,\zeta,b) & \mathfrak{R}_{11}(b,\zeta,w) & \mathfrak{R}_{12}(b,\zeta,w) \\ \mathfrak{T}_{21}(w,\zeta,b) & \mathfrak{T}_{22}(w,\zeta,b) & \mathfrak{R}_{21}(b,\zeta,w) & \mathfrak{R}_{22}(b,\zeta,w) \\ \mathfrak{R}_{11}(w,\zeta,b) & \mathfrak{R}_{12}(w,\zeta,b) & \mathfrak{T}_{11}(b,\zeta,w) & \mathfrak{T}_{12}(b,\zeta,w) \\ \mathfrak{R}_{21}(w,\zeta,b) & \mathfrak{R}_{22}(w,\zeta,b) & \mathfrak{T}_{21}(b,\zeta,w) & \mathfrak{T}_{22}(b,\zeta,w) \end{bmatrix} \begin{bmatrix} \hat{\underline{\underline{S}}}_1^+(w) \\ \hat{\underline{\underline{S}}}_2^+(w) \\ \hat{\underline{\underline{S}}}_1^-(b) \\ \hat{\underline{\underline{S}}}_2^-(b) \end{bmatrix} + \begin{bmatrix} \mathfrak{S}_1(w,\zeta,b) \\ \mathfrak{S}_2(w,\zeta,b) \\ \mathfrak{S}_1(b,\zeta,w) \\ \mathfrak{S}_2(b,\zeta,w) \end{bmatrix} \tag{5.92}$$

These equations are the invariant imbedding relationships for depths (w,ζ,b). The incident radiances at depth w and b are now known, so these equations allow the radiances at any depth ζ

in the water body $[w, m]$ to be computed, including all effects of the air-water surface, the bottom reflectance, and the internal sources. These equations are evaluated for each Fourier ℓ mode. It should be noted that the incident radiance at b is zero for an opaque reflecting bottom, so in practice only the $\mathfrak{T}_{ij}(w, \zeta, b)$ and $\mathfrak{R}_{ij}(w, \zeta, b)$ terms need to be computed. In this case we have, for example,

$$\underline{\hat{S}}_1^+(\zeta) = \mathfrak{T}_{11}(w, \zeta, b)\underline{\hat{S}}_1^+(w) + \mathfrak{T}_{12}(w, \zeta, b)\underline{\hat{S}}_1^+(w) + \mathfrak{S}_1(w, \zeta, b) \tag{5.93}$$

For the case of no polarization, $\mathfrak{T}_{12}(w, \zeta, b) = 0$, and this equation reduces to (the transpose of) L&W (8.103) with $p = 1$.

5.6 Computing the Water-leaving Radiances

The only radiances still unknown are the water-leaving radiances $\underline{\hat{S}}_{1,2}^-(a)$. Now that the upwelling radiances at depth w are known, the water-leaving radiances can be obtained from the surface boundary condition (5.18) (L&W 8.107):

$$\underline{\hat{S}}_p^-(a) = \underline{\hat{r}}_p(a, w)\underline{\hat{S}}_p^+(a) + \underline{\hat{t}}_p(w, a)\underline{\hat{S}}_p^-(w) \tag{5.94}$$

with $p = 1, 2$. Because the surface-bounday Fourier modes do not decouple, this equation must be evaluated simultaneously for all ℓ modes via the $4M(N+1) \times 4M(N+1)$ composite matrices as seen in Eqs. (5.15) and (5.16), and the $4M(N+1) \times 1$ vectors of Eq. (5.17).

5.7 Synthesis of the Physical-space Stokes Vectors

We have now obtained the Fourier amplitudes of the Stokes vectors at all depths. The final step of the solution of the VRTE is to reconstitute the quad averaged Stokes vectors in physical space. The $4M \times 1$ column vectors $\underline{\hat{S}}_{1,2}^{-\pm}$ are "unstacked" as in Eq. (4.23) to obtain the 4×1 Stokes vector amplitudes $\underline{\hat{S}}_{1,2}^{\pm}(\zeta, u|\ell)$ for each u and ℓ value. The physical Stokes vectors are then obtained via Eq. (4.3) (L&W 8.108):

$$\underline{S}^{\pm}(\zeta, u, v) = \sum_{\ell=0}^{N} \left[\underline{\hat{S}}_1^{\pm}(\zeta, u|\ell) \cos(\ell\phi_v) + \underline{\hat{S}}_2^{\pm}(\zeta, u|\ell) \sin(\ell\phi_v) \right] . \tag{5.95}$$

Evaluation of this equation at each depth $a, w = \zeta_1, \zeta_2, ..., \zeta_K = m$, for each value of $u = 1, ..., M$ and $v = 1, 2, ..., 2N$, gives the desired quad-averaged Stokes vectors.

This completes the solution of the vector radiative transfer equation.

CHAPTER 6

Flow Chart of the Computations

The preceding development of invariant imbedding theory for the VRTE has been, to put it mildly, rather complicated. The relevant equations are spread over almost 100 pages. It is therefore worthwhile to explicitly lay out the exact steps needed to implement the solution in the HydroPol software.

There are several distinct steps to the solution:

- **Set default values to be used for all subsequent VRTE solutions.** Certain quantities can be set to recommended default values. The main one of these is the quad partition.
- **The air-water transfer functions $\hat{\underline{\underline{r}}}_{1,2}(a,w)$, $\hat{\underline{\underline{r}}}_{1,2}(w,a)$, $\hat{\underline{\underline{t}}}_{1,2}(a,w)$, and $\hat{\underline{\underline{t}}}_{1,2}(w,a)$ must be computed.** These depend only on the quad partitioning, the water index of refraction, and the wind speed. For a given quad partition and index of refraction, they can therefore be computed as a function of the wind speed and stored in a library of surface functions for different wind speeds. The functions for the user requested wind speed can then be read and used in the solution for that particular problem.
- **The discretized phase matrices $\hat{\underline{P}}^{\pm}_{1,2}(\zeta,r,u|k)$ must be computed.** These depend only on the quad partition and the form of the phase matrix. A library of pre-computed phase matrices can be created so that the phase matrices requested by the user for different IOP components can be read and used for the particular problem at hand.
- **The user must specify the physical problem to be solved.** This includes defining the water-column IOPs, the depths and wavelengths where output is to be computed and saved, the incident sun and sky radiance distribution, the wind speed, and the water column bottom depth and reflectance.
- **The VRTE must be solved for the requested physical problem.** This requires carrying out all of the computations shown in Chapters 4 and 5 .

The fundamental output of the HydroPol software is a file of quad-averaged Stokes vector components $\underline{S} = [I,Q,U,V]^T$ as functions of depth z_k, direction $(u,v) = (\theta,\phi)$, and wavelength λ_j.

Derived quantities such as irradiances, reflectances, diffuse attenuation functions, and degree of polarization are computed from their definitions using the Stokes vector components. The user can then examine these outputs graphically or analytically outside of the HydroPol software.

The next sections lay out the needed computations step by step.

6.1 Specification of Default Values

These choices affect all subsequent calculations.

▶ **The quad partition.** The default quad partition shown in Fig. 2.1 has proven adequate for almost all problems in HydroLight, and should work equally well in HydroPol. This partition is set as the default.

▶ **The water index of refraction.** This is set to $n_{\text{water}} = 1.34$ for debugging. (HydroLight includes the option to have n_{water} depend on the user-selected salinity and temperature, but his requires a much larger table of surface transfer functions. That option can be added to HydroPol later if desired.)

▶ **Inelastic scatter functions.** The excitation-emission spectra (wavelength redistribution functions) for chlorophyll and CDOM fluorescence are set to the same values used in HydroLight. Likewise, the scattering cross section and phase matrix for Raman scatter by water are set to standard values. The details of the source terms have not been included in this report. In the first version of the code, inelastic scatter can be set to zero to simplify debugging.

These defaults can be changed by the user, but note that changing the quad resolution requires re-computation of the surface transfer function and phase matrix libraries, and changing the index of refraction requires re-computation of the surface transfer function library.

6.2 Computation of the Air-water Radiance Transfer Functions

The default quad partition and water index of refraction allow the air-water surface transfer functions to be computed for various wind speeds. These are one-time calculations for a given wind speed and index of refraction and are carried out by code that is separate from the main HydroPol code that solves the VRTE for a specific set of user inputs. For a level sea surface (a wind speed of 0), there are unique reflected and transmitted directions for a each incident direction, and the calculations can be based on the analytic Fresnel reflectance and transmittance functions of §3.1. For a rough sea surface (wind speed > 0), a given incident direction can be reflected and refracted by surface waves into any direction, and Monte Carlo calculations must be used to compute ensemble averages over many wave realizations.

6.2.1 Analytic Computation of Air-water Transfer Functions for a Level Surface

There are four transfer functions, $\underline{r}_{aw}$, $\underline{r}_{wa}$, $\underline{t}_{aw}$, and $\underline{r}_{wa}$, that must first be quad-averaged, and then Fourier analyzed. Each of these is a 4×4 matrix for each combination of incident (r, s) and reflected or refracted (u, v) quad directions. The process is the same for each element of each matrix. Let

$t(\theta_i, \phi_i, \theta_r, \phi_r)$ represent the i, j element of any of the four transfer matrices, with (θ_i, ϕ_i) being the incident direction and (θ_r, ϕ_r) being the reflected or refracted direction. For a level surface, the rotation matrices reduce to $\underline{I}_{4\times4}$ and thus do not need to be explicitly considered. The calculations proceed as follows.

▷ Begin loop over all incident quads $Q(r, s)$.

► For each $Q(r, s)$, the quad-averaged $t(r, s, u, v)$ defined by Eq. (2.11) is computed as a quadruple sum over subquads as shown in Eq. (2.15).

▷ Begin loop over subquads of $Q(r, s)$.

► For each subquad of $Q(r, s)$ in Eq. (2.15), compute the reflected or refracted direction using either the law of reflection ($\theta_r = \theta_i$) or Snel's law (3.1). This gives the angles needed to (1) determine which quad $Q(u, v)$ receives the reflected or transmitted light and (2) evaluate whichever matrix element is being computed, using Eqs. (3.2) to (3.23) as needed. Note that $\phi_r = \phi_i$ for a level surface. Note also that for reflection, each $Q(r, s)$ reflects light into exactly one $Q(u, v)$, with all other reflected quads having a value of zero. However, for transmission, $Q(r, s)$ can transmit light into more than one $Q(u, v)$. In either case, most of the matrix elements will be zero for a level surface because the quads are not connected by reflection or transmission.

► Compute the contribution to the summation of Eq. (2.15) using whichever of Eqs. (3.2) to (3.23) corresponds to the matrix element being computed.

◁ End of loop over subquads.
◁ End of loop over incident quads.

At the end of these computations, the quad-averaged $t(r, s \to u, v)$ are known for all quad pairs. These arrays are quad-averaged reflectances and transmittances in physical space, so they can be plotted or otherwise compared with the analytical functions as a check on the computer code.

► Compute the Fourier amplitudes $\hat{t}_1(r, u|k, \ell)$ and $\hat{t}_2(r, u|k, \ell)$ of each matrix element using Eqs. (5.7) and (5.8) and the special cases shown therein. The elements of the 4×4 Fourier amplitude matrices $\underline{\hat{t}}_{1,2}(r, u|k, \ell)$ are now known.

► Assemble the 4×4 $\underline{\hat{t}}_{1,2}(r, u|k, \ell)$ matrices into $4M \times 4M$ composite matrices $\underline{\hat{t}}_{1,2}(k, \ell)$ as shown in Eqs. (5.11) and (5.12).

► Assemble the $4M \times 4M$ composite matrices $\underline{\hat{t}}_{1,2}(k, \ell)$ into $4M(N+1) \times 4M(N+1)$ composite matrices $\underline{\hat{t}}_{1,2}$ as shown in Eqs. (5.15) and (5.16).

These calculations give the composite matrices $\underline{\hat{r}}_{1,2}(a, w)$, $\underline{\hat{r}}_{1,2}(w, a)$, $\underline{\hat{t}}_{1,2}(a, w)$, and $\underline{\hat{t}}_{1,2}(w, a)$, which appear in the surface boundary conditions (5.18) and (5.19).

► Store these matrices in a file identified by the wind speed and water index of refraction, e.g. with a file name like `SurfrtU0n134.txt`.

This completes the computation of the surface transfer functions for a level sea surface.

6.2.2 Monte Carlo Computation of Air-water Transfer Functions for a Wind-blown Surface

When the sea surface is not level, Monte Carlo simulation for an ensemble of randomly generated surface waves is the only way to compute surface transfer functions that include multiple scattering effects for arbitrary wave surfaces. The details of those ray tracing calculations are described in §3.2 and in Appendix B, and need not be repeated here. The essence of these calculations can be summarized as follows:

▷ Begin loop over all surface realizations.

▷ Begin loop over all incident quads $Q(r, s)$.

► For each $Q(r, s)$, choose an initial ray direction at random from within the solid angle subtended by the quad.

► Trace the initial ray and all of its daughter rays to completion. Whenever a ray leaves the hexagonal domain, tally the ray's final weight array $\underline{\mathcal{W}}$ to one of the four accumulating weight arrays, as determined by the ray's initial and final quads.

◁ End of loop over incident quads.
◁ End of loop over surface realizations.

► Divide the accumulating weight arrays by the number of surface realizations to obtain the ensemble averages of Eq. (3.56). Convert the ensemble averages to the four physical space, quad-averaged transfer functions $t(r, s \rightarrow u, v)$ via Eq. (3.67).

At the end of these ray-tracing calculations, the four transfer functions $t(r, s \rightarrow u, v)$ correspond to those obtained for the level surface case (but of course have different numerical values). The Fourier decomposition then proceeds in exactly the manner described above for the level surface transfer functions.

6.3 Computation of the Normalized Phase Matrices

Continuous-variable phase matrices must be directionally discretized and converted to Fourier amplitude form. This is most conveniently done by factoring out the scattering coefficient as shown in Eq. (1.27) and working with the normalized phase matrices $\underline{\tilde{P}}$. The normalized phase matrices are in turn most conveniently (for the user) defined by separately selecting a scattering phase function $\tilde{\beta}$ (the 1,1 element of the normalized phase matrix) and a reduced scattering matrix $\underline{\tilde{M}}$, as shown in Eq. (1.25): $\underline{\tilde{P}} = \tilde{\beta}\,\underline{R}\,\underline{\tilde{M}}\,\underline{R}$, where $\underline{R}$ is a rotation matrix as seen in Eq.(1.31). Unfortunately it is not possible to separately discretize phase functions and reduced scattering matrices or reduced phase matrices (as seen in Eq. 1.26), building up a library of each, and then to combine them at run time according to a user's choices for a particular problem. The reason is simply that the discretization process involves integrals, and the integral of a product is not the product of the integrals. Thus each combination of a $\underline{\tilde{P}}$ and an $\underline{\tilde{M}}$ (including the rotation matrices) must be discretized separately. However, in practice this should not a great inconvenience. Although a wide

range of phase functions $\tilde{\beta}$ is needed for ocean applications, these can be paired with relatively few $\underline{\tilde{M}}$ matrices. As noted in §1.4.1, only a Rayleigh reduced scattering matrix, $\underline{\tilde{M}}_{\text{Ray}}$, and perhaps one or two more based on measurements, need to be considered. Thus we can create a library of many phase functions paired with just one or a few reduced phase matrices, e.g. as seen in Eq. (1.34). In principle, both $\tilde{\beta}$ and $\underline{\tilde{M}}$ are functions of depth and wavelength, but in practice the depth and wavelength dependence can to a reasonable approximation be placed in the scattering coefficient $b(z, \lambda)$ (as is done in HydroLight for for volume scttering functions). $\tilde{\beta}$ and $\underline{\tilde{M}}$ are then functions only of the scattering angle ψ.

These are one-time calculations for a phase function and reduced phase matrix and are carried out by code that is separate from the main HydroPol code that solves the VRTE for a specific set of user inputs. The calculations for discretizing and Fourier decomposing normalized phase matrices are as follows.

▶ Select a scattering phase function $\tilde{\beta}(\psi)$ and a reduced scattering matrix $\underline{\tilde{M}}(\psi)$.

▷ Begin loop over all pairs of incident $Q(r, s)$ and scattered $Q(u, v)$ quads.
Separate calculations are done for the $\underline{\tilde{P}}^{+}$ and $\underline{\tilde{P}}^{-}$ matrices as seen in Eq. (4.2).

For a given pair of $Q(r, s)$ and $Q(u, v)$, the quad-averaged normalized phase matrix is computed as a quadruple sum as in Eq. (2.15), where F in that equation is now $\tilde{\beta}\,\underline{R}\,\underline{\tilde{M}}\,\underline{R}$.

▷ Begin loop over all pairs of subquads in $Q(r, s)$ and $Q(u, v)$, as described in §2.2.1.

▶ For each subquad of $Q(r, s)$ and subquad of $Q(u, v)$ in Eq. (2.15), compute the scattering angle ψ by Eq. (2.16).

▶ For each subquad of $Q(r, s)$ and subquad of $Q(u, v)$, compute the rotation angles α and α' by Eqs. (1.5-1.12), using the (θ', ϕ') of the $Q(r, s)$ subquad and the (θ, ϕ) of the $Q(u, v)$ subquad and the ψ value just computed. Note that the scattering and rotation angles are different for the same (r, s, u, v) values, depending on whether (r, s) and (u, v) are in the same or different hemispheres, as shown in Fig. 4.2.

▶ Compute the 16 elements of the normalized phase matrix $\underline{\tilde{P}}$ for this pair of subquads as shown in Eq. (1.32).

▶ Compute the contribution of the this pair of subquads to the summations in Eq. (2.15) for the given $Q(r, s)$ and $Q(u, v)$.

◁ End loop over subquad pairs.
◁ End loop over quad pairs.

▶ After computing all of the quad averages, use Eq. (2.19) to recompute the forward-scatter quads $\underline{\tilde{P}}^{\pm}_{1,1}$ to correct for any numerical error in those calculations.

The quad-averaged normalized phase matrix $\underline{\tilde{P}}^{\pm}(r, s \rightarrow u, v)$ is now known for all pairs of incident and scattered quads.

▶ Compute the Fourier amplitudes $\underline{\hat{P}}^{\pm}_{1,2}(r,u|k)$ of the $\underline{\tilde{P}}^{\pm}(r,s\rightarrow u,v)$ using Eqs. (A.12) and (A.14). Note that, as explained in §1.4.2 and illustrated in Fig. 1.5, the upper-left and lower-right 2×2 blocks of the phase matrix can be expanded as a cosine series; the $\underline{\hat{P}}^{\pm}_{2}$ sine amplitudes for those elements can be set to 0. Similarly, the upper-right and lower-left 2×2 blocks of the phase matrix can be expanded as a sine series; the $\underline{\hat{P}}^{\pm}_{1}$ cosine amplitudes for those elements can be set to 0.

The 4×4 amplitude matrices $\underline{\hat{P}}^{\pm}_{1,2}$ are now known.

▶ Store these matrices in a file identified by the phase function and reduced scattering matrix, e.g. with a file name like `FFpf02pmRay.txt` for a Fournier-Forand phase function with a backscatter fraction of 0.02 and a Rayleigh reduced phase matrix.

This completes the computation of the normalized phase matrix Fourier amplitudes for the chosen $\tilde{\beta}$ and $\underline{\tilde{M}}$.

6.4 User Specification of the Physical Problem

It is emphasized that HydroPol (like HydroLight) simply solves the radiative transfer equation for whatever conditions the user requests. The solution equations *per se* are valid for any medium and problem for which radiative transfer theory is valid, and the software does not inherently know anything about the ocean. *It is therefore up to the user to supply all of the information needed to define the oceanic (or other) problem of interest.* In HydroLight, the information needed to solve the SRTE is solicited and organized via a graphical user interface, and there are many options for how the inputs can be defined (e.g., via bio-geo-optical models that convert concentrations to absorption and scattering properties). However, use of that interface is not required: all it does is create a file with the values of various inputs, flags, file names, and the like, which are then read by HydroLight. The same scheme is used by HydroPol. The only requirement is that the user must define, one way or another, the following inputs:

▶ **The depths where the solution is to be saved.** The user inputs a list of depths $z_i, i=1,...,K$, with $z_1 = w = 0$ being depth 0 just below the mean air-water surface, and with $z_K = m$ being the greatest depth where output is to be saved. Output is also saved at "depth" a, in the air just above the water surface.

▶ **The wavelengths where the solution is to be saved.** The user inputs either a single wavelength or a list of wavelength *band boundaries*, $\lambda_j, j=1,...,J$. In the latter case, the IOPs are evaluated and the output will be saved at the centers of the bands. Thus if $J>1$, the first nominal output wavelength is $(\lambda_1+\lambda_2)/2$, and so on.

▶ **The albedo of single scattering $\omega_o(\zeta,\lambda)$.** In the oceanographic setting, it is usually convenient for the user to specify the total absorption coefficient $a(z,\lambda)$ and the total scattering coefficient $b(z,\lambda)$. IOPs are additive, so these are in turn usually computed as sums of contributions by pure water and whatever particles (phytoplankton, minerals, bubbles, etc.) and dissolved substances (colored dissolved organic matter, pollutants, etc.) are in the water column. HydroPol

then internally computes the beam attenuation, $c(z,\lambda) = a(z,\lambda) + b(z,\lambda)$; uses $c(z,\lambda)$ in Eq. (1.20) to compute the correspondence between optical depth ζ and physical depth z; and computes $\omega_o(\zeta,\lambda) = b(z,\lambda)/c(z,\lambda)$. However, in idealized simulations such as homogeneous water, it is often convenient to specify ω_o directly.

► **The normalized phase matrix $\underline{\tilde{P}}$.** As with absorption and scattering coefficients, the total phase matrix $\underline{P}$ is the sum of the phase matrices for whatever scattering particles are in the water. As shown in Eq. (1.27), it is common to factor $\underline{P}$ into the product of the scattering coefficient and the normalized phase matrix, $\underline{P} = b\,\underline{\tilde{P}}$. The total normalized phase matrix can therefore be computed as a scattering-weighted sum of normalized phase matrices for the various water components:

$$\underline{\tilde{P}} = \sum_i \left(\frac{b_i}{b}\right) \underline{\tilde{P}}_i \,, \tag{6.1}$$

where the sum is over all of the scattering components in the water. The user can specify the total $\underline{\tilde{P}}$, or can select phase matrices $\underline{\tilde{P}}_i$ for each water component (phytoplankton, mineral, etc.), in which case HydroPol will compute the needed total using Eq. (6.1). Phase matrices are functions of incident and scattered directions, including the effects of the rotation matrices. In practice, the phase matrices are usually specified via the choice of a scattering phase function $\tilde{\beta}(\psi)$ and a reduced scattering matrix $\underline{\tilde{M}}(\psi)$, as shown in Eq. (1.25) and discussed in §6.3. The effects of the rotation angles are then incorporated by HydroPol during quad discretiation to create the quad-averaged phase matrices for the chosen quad partition.

For initial debugging, the user can select one of the previously processed normalized phase matrices from the library of such files for the $\underline{\tilde{P}}_i$. Later on, as in HydroLight, the user can request values with a particular backscatter fraction, in which case the needed $\underline{\tilde{P}}_i$ can be obtained by interpolation between the available values.

The IOPs must be defined at all wavelengths $\lambda_1 \le \lambda \le \lambda_J$ and all depths $z_1 = 0 \le z \le z_k = m$, not just at the discrete values where output is requested. This is because the ODE solver takes very small and variable-sized steps in optical depth and may thus request IOP values at any depth between 0 and m. However, it is common to define the IOPs only at discrete depths and wavelengths, in which case HydroPol will interpolate between the given values as needed to obtain values at other depths and wavelengths. *The discrete depths and wavelengths where the IOPs are defined do not need to be the same discrete values where output is to be saved.*

► **The wind speed.** The wind speed is needed to model the surface roughness. The reflectance and transmission properties of both level and wind-blown surfaces are pre-computed as described in Chapter 3, Fourier analyzed, and stored in a library. At run time, the surface transfer function Fourier amplitudes for the requested wind speed are read from the library of surface transfer functions.

► **The incident sky radiance.** The Stokes vector $\underline{S}^+(a,\theta,\phi,\lambda)$ just above the sea surface must be defined for all downwelling directions and all wavelengths $\lambda_1 \le \lambda \le \lambda_J$. This vector is the total radiance of the sun's direct beam and the background sky. $\underline{S}^+(a,\theta,\phi,\lambda)$ can be obtained from measurements read from a datafile, or from ancillary models. Because HydroPol does not (yet, at

least) have an associated atmospheric vector radiative transfer model, the default is to use the same atmospheric sky radiance models as in HydroLight. These provide only unpolarized sun and sky radiance. The actual sky radiance is partially polarized, so an unpolarized input sky radiance is not correct. However, an unpolarized sky radiance will be adequate for initial debugging.

The unpolarized HydroLight sky radiance models require various kinds of atmospheric inputs including aerosol type, water vapor, sea-level pressure, cloud fraction, wind speed, etc., as well as the zolar zenith angle. These inputs can be given in various ways, including specification of latitude, longitude, day of year, and universal time, from which the solar zenith angle is computed by HydroPol.

► **The bottom reflectance.** As was seen in Eq. (5.2), the opaque bottom boundary at depth $z_K = m$ is assumed to azimuthally isotropic with a known vector bidirectional reflectance distribution function of the form $\underline{VBRDF}(\theta', \theta, \phi - \phi', \lambda)$. This function must be defined for any incident and reflected directions, and for all wavelength in the requested range. As noted, for initial debugging, and often in practice, it will be sufficient to assume that the bottom is a Lambertain reflector. In that case, only the irradiance reflectance of the bottom, $R(\lambda)$ needs to be specified by the user.

► **The internal sources.** If there are internal sources of bioluminescence, chlorophyll or CDOM fluorescence, or Raman scatter by the water, these must be specified via appropriate source terms $\underline{\Sigma}(\zeta, \mu, \phi, \lambda)$ as seen in Eq. (1.22). Bioluminescence is assumed to emit unpolarized light, as do chlorophyll and CDOM fluorescence. In practice, default functions are used to convert chlorophyll or CDOM concentrations into the appropriate source terms. Raman scatter by water does depend on polarization, but this has a fixed mathematical form. Thus all the user needs to do is decided whether or not to include Raman scatter in the run. For initial debugging, internal sources will be omitted. Details of the mathematical formulations of the various kinds of source terms $\underline{\Sigma}(\zeta, \mu, \phi, \lambda)$ will be added after the code is fully debugged for inelastic scatter.

► **Set various flags.** These flags can specify what output files are to be created, e.g., full output of all Stokes vector components, selected output for opening as Excel® spreadsheets, or output files formatted for reading by IDL®, MATLAB® or other graphics routines.

6.5 Solution of the VRTE

The information needed to solve the VRTE for a specific problem is now available. These calculations are carried out by the main HydroPol code.

► Initialize the run. This includes reading in the user's inputs, which define the particular problem to be solved, as well as reading in the pre-computed files of surface transfer functions and phase matrices that will be needed.

▹ Begin loop over wavelength.

► Read in or compute the incident sky radiance distribution $\underline{S}^+(a, \theta, \phi)$ for the current wavelength.

▶ Compute the quad-average sky radiances $\underline{S}^+(a,r,s)$ as a double summation using Eq. (2.14). This calculation is done independently for each component $[I,Q,U,V]^T$ of the 4×1 Stokes vector $\underline{S}^+(a,r,s)$.

▶ Compute the Fourier cosine $\hat{\underline{S}}_1^+(a,u|\ell)$ and sine $\hat{\underline{S}}_2^+(a,u|\ell)$ amplitudes of the discretized sky radiances using Eqs. (A.7) and (A.9).

▶ Store the 4×1 amplitude vectors as $4M\times 1$ composite column vectors $\hat{\underline{\underline{S}}}_{1,2}^+(a|\ell)$ as in Eq. (4.23).

▶ If internal sources are included, quad average and Fourier decompose each of the needed source terms $\underline{\Sigma}(\theta,\phi)$ in the same manner as was just shown for the sky radiance. The end result will be both upward and downward composite vectors $\hat{\underline{\underline{\Sigma}}}^{\pm}(\ell)$ as seen in Eqs. (4.25) to (4.28).

▶Quad average and Fourier decompose the bottom boundary reflectance $r_{mb}(\theta',\phi'\rightarrow\theta,\phi)$, which is seen in Eq. (1.43). In general, this is done the same as for any function of two directions. However, for azimuthally isotropic bottom boundaries, $r_{mb}(\theta',\theta,\phi-\phi')$, as assumed in HydroPol, the quad discretization can proceed as before, but only a Fourier cosine expansion is needed. Compute the cosine amplitudes $\hat{\underline{r}}_1(m,b|r,u|k)$ using Eq. (5.28).

For the case of a Lambertian depolarizing bottom as defined in Eq. (1.45), the results are particularly simple and are given by Eq. (5.40).

Everything is now ready for the solution of the ODEs (4.40) to (4.70) for the standard matrices in the interior of the "bare slab" water body $[w,m]$. This solution is performed independently for each Fourier mode, $\ell=0,1,...,N$. In HydroLight the ODE solutions typically take about 90% of the total run time, and the same will surely be true in HydroPol. Therefore care must be used to make this part of the code as efficient as possible.

▷ Begin loop over Fourier modes, $\ell=0,1,...,N$.

▶ Begin simultaneous integration of the downward sweep equations (4.40)-(4.44) starting with initial conditions (4.45)-(4.47) and of Eqns. (4.48)-(4.54) starting with (4.55)-(4.58). These compute $\hat{\underline{\underline{R}}}_1(\zeta,w)$, ..., $\hat{\underline{\underline{T}}}_1(\zeta,w)$ and $\hat{\underline{\underline{\Sigma}}}_{1,2}^{+t}(w,\zeta)$.

▶ During the ODE integrations, define the 4×4 elements of the $4M\times 4M$ local reflectance $\hat{\underline{\underline{\rho}}}_1(\zeta|\ell)$ and transmittance $\hat{\underline{\underline{\tau}}}_1(\zeta|\ell)$ matrices using Eqs. (4.15a) to (4.18f). The quantities depend on the IOPs at the current depth and wavelength via ω_o and $\underline{P}_{1,2}^{\pm}$. The composite matrices $\hat{\underline{\underline{\rho}}}_1(\zeta|\ell)$ and $\hat{\underline{\underline{\tau}}}_1(\zeta|\ell)$ are ordered as in Eq. (4.24).

During the ODE integration, save the results at each user-requested output depth. That is, save the matrices $\hat{\underline{\underline{R}}}_1(\zeta_1=w=0,w|\ell)=\underline{\underline{0}}_{4M\times 4M}$, $\hat{\underline{\underline{R}}}_1(\zeta_2,w|\ell)$, ..., $\hat{\underline{\underline{R}}}_1(\zeta_K=m,w|\ell)$, and so on for all of the standard operators.

▶ Begin simultaneous integration of the upward sweep Eqns. (4.59)-(4.70) starting with initial conditions (5.36)-(5.39) for the user-selected bottom boundary. This integration gives $\hat{\underline{\underline{R}}}_1(\zeta,m)$,

..., $\hat{\underline{\underline{T}}}_2(\zeta, m)$ and $\hat{\underline{\underline{\Sigma}}}_{1,2}^{+t}(\zeta, m)$. As for the downward sweep, save the matrices at each user-requested depth.

Note that although the ODEs are most conveniently written as matrix equations, the ODE solver requires one long vector of equations, rather than several individual matrix equations. The matrix equations therefore need to be "unpacked" into a single vector as shown in Eq. (4.74) before passing to the ODE solver.

Note also that the upward and downward sweep integrations are independent and time consuming (in HydroLight, solving the scalar versions of these equations accounts for about 90% of the run time). Integration of these equations is thus a good place to parallelize the code, which potentially could cut the wall-clock run time in half.

◁ End loop over ℓ for the ODE integrations. The standard operators for the "bare-surface" water body plus the bottom, $[w, b]$, are now known.

▶ Combine the $4M \times 4M$ composite matrices for the standard operators for each ℓ mode into $4M(N+1) \times 4M(N+1)$ composite block-diagonal matrices $\hat{\underline{\underline{R}}}_1(\zeta, w)$ etc. as shown in Eq. (5.20).

The next calculations incorporate the air-water surface $[a, w]$ with the water body $[w, \zeta]$. These calculations must be done for all ℓ modes simultaneously.

▶ Compute the $4M(N+1) \times 4M(N+1)$ complete operators $\mathfrak{T}_{11}(a, w, b)$ to $\mathfrak{R}_{22}(a, w, b)$ seen in Eq. (5.75). Note that only the 8 (a, w, b) matrices in the left two columns of Eq. (5.75) are needed because the upward radiance at the bottom is 0. The (b, w, a) matrices do not need to be computed.

$\mathfrak{T}_{11}(a, w, b)$ etc. are computed by equations of the form of equation set (5.49), with $r_p = \hat{\underline{\underline{r}}}_p(w, a)$, $t_p = \hat{\underline{t}}_p(a, w)$ as before, but now with $R_1 = \hat{\underline{\underline{R}}}_1(w, b)$,..., $T_2 = \hat{\underline{\underline{T}}}_2(b, w)$, and $\Sigma_p = \hat{\underline{\underline{\Sigma}}}_p^{-t}(b, w)$. This substitution is illustrated in Eq. (5.80). We can now set $\sigma_p = \hat{\underline{\underline{\sigma}}}_p^{+t}(a, w) = 0$ because the air-water surface is now taken to be source free.

▶ Compute the radiances at depth w using Eqs. (5.76)-(5.79).

The next calculations compute the radiances at all user requested depths ζ using the radiances at depth w. These calculations are done ℓ mode by ℓ mode.

▷ Begin loop over Fourier modes, $\ell = 0, 1, ..., N$.

▷ Begin loop over depths, $\zeta = \zeta_1 = 0, \zeta_2, ..., \zeta_K = m$.

▶ Compute the $4M(N+1) \times 4M(N+1)$ complete operators $\mathfrak{T}_{11}(w, \zeta, b)$ to $\mathfrak{R}_{22}(w, \zeta, b)$ seen in Eq. (5.92). Note that only the 8 (w, ζ, b) matrices in the left two columns of Eq. (5.92) are needed because the upward radiance at the bottom is 0. The (b, ζ, a) matrices do not need to be computed.

The $\mathfrak{T}_{11}(w, \zeta, b)$ are computed as shown in Eq. (5.91) using the n_{ij} and p_{ij} defined in Eqs. (5.86) and (5.88), respectively.

▶ Compute the radiances at each depth ζ and ℓ value using Eq. (5.92), as illustrated in Eq. (5.93).

◁ End loop over depths.

◁ End loop over Fourier modes.

▶ Compute the water-leaving radiances using Eq. (5.94). This calculation must be done for all ℓ modes at once using the $4M(N+1) \times 4M(N+1)$ composite matrices and $4M(N+1) \times 1$ composite Stokes vector amplitudes.

The amplitudes of the Stokes vectors are now known at all depths $a, \zeta_1 = w, \zeta_2, ..., \zeta_K = m$. The final step is to reconstitute the physical Stokes vectors from their Fourier amplitudes.

▶ Compute the physical Stokes vectors using Eq. (5.95).

This completes the solution of the VRTE for the Stokes vectors at all depths and directions, at the current wavelength.

▶ Compute derived quantities as desired, e.g. irradiances and remote-sensing reflectance.

▶ Write output files for the current wavelength as desired. Note that if fluorescence is included in the run, the scalar irradiance must be saved at each wavelength for use in generating (isotropic, unpolarized) the internal sources at subsequent wavelengths. If Raman scatter is included, the full radiance distribution must be saved for generation of the (non-isotropic, polarization-dependent) Raman source term at subsequent wavelengths.

◁ End loop over wavelengths.

▶ Compute quantities that depend on all wavelengths, e.g., Secchi depth and chromaticity coordinates.

▶ Finalize the run: e.g., close output files and report the total run time.

APPENDIX A

Fourier Analysis of Discrete Functions

The invariant imbedding solution algorithm for the directionally discretized VRTE employs a Fourier decomposition of the equations in azimuthal direction. This appendix collects for reference various results for the Fourier decomposition of discrete functions of azimuthal angle, as needed in Chapter 4. Most of this material comes from *Light and Water*, §8.3.

A.1 Functions of one azimuthal angle

Let $f_v \equiv f(\phi_v)$ be any discrete function of the azimuthal angle ϕ, i.e., f_v is defined only at the discrete values $\phi_v = (v-1)\Delta\phi$, for $v = 1, 2, ..., 2N$. Here $\Delta\phi = 2\pi/2N$ is the angular width of a quad as shown in Fig. 2.1; $\Delta\phi$ is the same for each quad. Let δ_k be the Kronecker delta function defined by

$$\delta_k \equiv \begin{cases} 1 & \text{if } k = 0 \\ 0 & \text{if } k \neq 0 \end{cases} \tag{A.1}$$

or equivalently

$$\delta_{k-\ell} \equiv \begin{cases} 1 & \text{if } k = \ell \\ 0 & \text{if } k \neq \ell \end{cases} \tag{A.2}$$

where k and ℓ are integers.

The discrete orthogonality relations for sines and cosines can then be written

$$\begin{aligned} \sum_{s=1}^{2N} \cos(k\phi_s)\cos(\ell\phi_s) &= \begin{cases} 0 & \text{if } k \neq \ell \\ N & \text{if } k = \ell \text{ and } \ell = 1, 2, ..., N-1 \\ 2N & \text{if } k = \ell \text{ and } \ell = 0 \text{ or } \ell = N \end{cases} \\ &= N(\delta_{k+\ell} + \delta_{k-\ell} + \delta_{k+\ell-2N}) \end{aligned} \tag{A.3}$$

$$\sum_{s=1}^{2N} \sin(k\phi_s)\sin(\ell\phi_s) = \begin{cases} 0 & \text{if } k \neq \ell \\ N & \text{if } k = \ell \text{ and } \ell = 1, 2, ..., N-1 \\ 0 & \text{if } k = \ell \text{ and } \ell = 0 \text{ or } \ell = N \end{cases}$$
$$= N(\delta_{k-\ell} - \delta_{k+\ell} - \delta_{k+\ell-2N}) \tag{A.4}$$

and

$$\sum_{s=1}^{2N} \cos(k\phi_s)\sin(\ell\phi_s) = 0 \quad \text{for all } k \text{ and } \ell\ . \tag{A.5}$$

A discrete function of one azimuthal angle (e.g., a discretized Stokes vector $\underline{S}(\zeta, u, v, j)$ with the depth ζ, polar angle u, and wavelength j being held constant) has the Fourier spectral decomposition

$$f(v) = \sum_{\ell=0}^{N} \left[\hat{f}_1(\ell)\cos(\ell\phi_v) + \hat{f}_2(\ell)\sin(\ell\phi_v)\right] \tag{A.6}$$

where $v = 1, 2, ..., 2N$. The notation $\hat{f}$ denotes a Fourier amplitude of the corresponding physical variable f; subscript 1 denotes cosine amplitudes and subscript 2 denotes sine amplitudes. The cosine amplitudes $\hat{f}_1(\ell)$ are obtained by multiplying Eq. (A.6) by $\cos(k\phi_v)$, summing over v, and applying the orthogonality relations to get

$$\hat{f}_1(\ell) = \frac{1}{\epsilon_l} \sum_{v=1}^{2N} f(v)\cos(\ell\phi_v) \qquad \text{for } \ell = 0, 1, ..., N \tag{A.7}$$

where

$$\epsilon_\ell = N(1 + \delta_{2\ell} + \delta_{2\ell-2N}) = \begin{cases} 2N & \text{if } \ell = 0 \text{ or } \ell = N \\ N & \text{if } \ell = 1, 2, ..., N-1 \end{cases} \tag{A.8}$$

Note that the $\hat{f}_1(0)$ amplitude is just the average value of $f(v)$. The sine amplitudes $\hat{f}_2(\ell)$ are obtained in a similar way by multiplying (A.6) by $\sin(k\phi_v)$:

$$\hat{f}_2(\ell) = \frac{1}{\gamma_l} \sum_{v=1}^{2N} f(v)\cos(\ell\phi_v) \qquad \text{for } \ell = 1, ..., N-1 \tag{A.9}$$

where

$$\gamma_\ell = N(1 - \delta_{2\ell} - \delta_{2\ell-2N}) = \begin{cases} 0 & \text{if } \ell = 0 \text{ or } \ell = N \\ N & \text{if } \ell = 1, 2, ..., N-1 \end{cases} \tag{A.10}$$

Note that values of $\gamma_0 = \gamma_N = 0$ do not occur in Eq. (A.9) for the sine amplitudes.

If $f(v)$ is a constant f then $\hat{f}_1(0) = f$ and all other cosine and sine components are zero. This is the case for a polar cap radiance, which has no azimuthal dependence. In general, the $2N$ values of $f(v)$ are determined exactly by the $N + 1$ cosine amplitudes and the $N - 1$ sine amplitudes; the information content of the physical and Fourier representations is the same. In the Fourier decomposition of discrete functions all terms must be included; there can be no truncation of these summations.

A.2 Functions of the difference of two azimuthal angles

Let $g_{\cos}(v,s) = g[\cos(\phi_v - \phi_s)]$ be a discrete cosine function of $\phi_v - \phi_s$, where ϕ_v and ϕ_s are two azimuthal angles and $v, s = 1, ..., 2N$. Then $g_{\cos}(v,s)$ has the Fourier expansion (L&W 8.24)

$$g_{\cos}(v,s) = \sum_{k=0}^{N} \hat{g}_1(k) \cos[k(\phi_v - \phi_s)] \,. \tag{A.11}$$

Multiplying this equation by $\cos(\ell\phi_v)$, summing over v, expanding the cosine, and using the orthogonality relations gives (L&W 8.25)

$$\hat{g}_1(\ell) = \frac{1}{\epsilon_\ell \cos(\ell\phi_s)} \sum_{v=1}^{2N} g_{\cos}(v,s) \cos \ell\phi_v) \qquad \text{for } \ell = 0, ..., N \,. \tag{A.12}$$

Similarly, let $g_{\sin}(v,s) = g[\sin(\phi_v - \phi_s)]$ be a discrete sine function of $\phi_v - \phi_s$. Then $g_{\sin}(v,s)$ has the Fourier expansion

$$g_{\sin}(v,s) = \sum_{k=1}^{N-1} \hat{g}_2(k) \sin[k(\phi_v - \phi_s)] \,. \tag{A.13}$$

Multiplying this equation by $\sin(\ell\phi_v)$, summing over v, expanding the sine, and using the orthogonality relations gives

$$\hat{g}_2(\ell) = \frac{1}{\gamma_\ell \cos(\ell\phi_s)} \sum_{v=1}^{2N} g_{\sin}(v,s) \sin \ell\phi_v) \qquad \text{for } \ell = 1, ..., N-1 \,. \tag{A.14}$$

These expansions will be used for the elements of the phase matrix, which depend on either the cosine or the sine of $\phi_v - \phi_s$. Note that in Eqs. (A.12) and (A.14) we can without loss of generality set $\phi_s = \phi_1 = 0$. This merely anchors the difference $\phi_v - \phi_s$ to $\phi_1 = 0$. The computer code then needs to compute and store phase function elements only for the range of $v = 1, ..., N+1$, which generates all discretized scattering angles ψ as $\phi_v = 0$ to 180 deg. The phase matrix element $\underline{P}(r, s \to u, v)$ is then obtained from the stored value of $\underline{P}(r, 1 \to u, v - s + 1)$, which can be stored as a three-index array $\underline{P}(r, u, v)$.

A.3 Functions of two azimuthal angles

Finally, let $h(s,v) = h(\phi_s, \phi_v)$ be an arbitrary function of two azimuthal angles. Then $h(s,v)$ can be represented as

$$\begin{aligned} h(s,v) = &\sum_{k=0}^{N}\sum_{\ell=0}^{N} \hat{h}_{11}(k,\ell) \cos(k\phi_s) \cos(\ell\phi_v) \\ &+ \sum_{k=0}^{N}\sum_{\ell=0}^{N} \hat{h}_{12}(k,\ell) \cos(k\phi_s) \sin(\ell\phi_v) \\ &+ \sum_{k=0}^{N}\sum_{\ell=0}^{N} \hat{h}_{21}(k,\ell) \sin(k\phi_s) \cos(\ell\phi_v) \\ &+ \sum_{k=0}^{N}\sum_{\ell=0}^{N} \hat{h}_{22}(k,\ell) \sin(k\phi_s) \sin(\ell\phi_v) \,. \end{aligned} \tag{A.15}$$

To find $\hat{h}_{11}$, multiply Eq. (A.15) by $\cos(k'\phi_s)\cos(\ell'\phi_v)$, sum over s and v, and apply the orthogonality relations. The other three amplitudes are found in an analogous manner. The results are

$$\begin{aligned}
\hat{h}_{11}(k,\ell) &= \frac{1}{\epsilon_k\epsilon_\ell}\sum_{s=1}^{2N}\sum_{v=1}^{2N} h(s,v)\cos(k\phi_s)\cos(\ell\phi_v)\,,\\
\hat{h}_{12}(k,\ell) &= \frac{1}{\epsilon_k\gamma_\ell}\sum_{s=1}^{2N}\sum_{v=1}^{2N} h(s,v)\cos(k\phi_s)\sin(\ell\phi_v)\,,\\
\hat{h}_{21}(k,\ell) &= \frac{1}{\gamma_k\epsilon_\ell}\sum_{s=1}^{2N}\sum_{v=1}^{2N} h(s,v)\sin(k\phi_s)\cos(\ell\phi_v)\,,\\
\hat{h}_{22}(k,\ell) &= \frac{1}{\gamma_k\gamma_\ell}\sum_{s=1}^{2N}\sum_{v=1}^{2N} h(s,v)\sin(k\phi_s)\sin(\ell\phi_v)\,.
\end{aligned} \tag{A.16}$$

The arbitrary zero sine amplitudes $\hat{f}_2(0)$ and $\hat{f}_2(N)$ have their counterparts here:

$$\begin{aligned}
&\hat{h}_{12}(k,0) = \hat{h}_{12}(k,N) = 0 \qquad \text{for k} = 0,...,\text{N}\,,\\
&\hat{h}_{21}(0,\ell) = \hat{h}_{21}(N,\ell) = 0 \qquad \text{for } \ell = 0,...,\text{N}\,,\\
&\hat{h}_{22}(0,0) = \hat{h}_{22}(0,N) = \hat{h}_{22}(N,0) = \hat{h}_{22}(N,N) = 0\,.
\end{aligned} \tag{A.17}$$

These special cases allow the exclusion of any k or ℓ values in Eq. (A.16) that would result in division by zero resulting from the γ_k and γ_ℓ factors.

Equations (A.15) to (A.17) will be applied to the air-water surface transfer functions. However, the symmetries of those functions will result in considerable simplification. In particular, all $\hat{h}_{12}(k,\ell)$ and $\hat{h}_{21}(k,\ell)$ will turn out to be zero.

The equations of Chapter 3 contain sums over k and ℓ having the form

$$S1 = \sum_{l=0}^{N}\sum_{k=0}^{N} f(k)g(\ell)\left[\sum_{s=1}^{2N}\cos[k(\phi_v-\phi_s)]\cos(\ell\phi_s)\right] \tag{A.18}$$

where $f(k)$ and $g(\ell)$ are discrete functions of k and ℓ for $k,\ell = 0,...,N$. Application of the the trigonometric formula for the cosine of the difference of two angles and the previous equations reduces this to

$$\begin{aligned}
S1 &= \sum_{l=0}^{N}\sum_{k=0}^{N} f(k)g(\ell)\ \left\{\left[\sum_{s=1}^{2N}\cos(k\phi_s)\cos(\ell\phi_s)\right]\cos(k\phi_v) + \left[\sum_{s=1}^{2N}\sin(k\phi_s)\cos(\ell\phi_s)\right]\sin(k\phi_v)\right\}\\
&= \sum_{l=0}^{N}\sum_{k=0}^{N} f(k)g(\ell)\ N(\delta_{k+\ell}+\delta_{k-\ell}+\delta_{k+\ell-2N})\cos(k\phi_v)\\
&= \sum_{l=0}^{N} f(\ell)g(\ell)\ N(1+\delta_{2\ell}+\delta_{2\ell-2N})\cos(\ell\phi_v)\\
&= \sum_{l=0}^{N} \epsilon_\ell f(\ell)g(\ell)\cos(\ell\phi_v)
\end{aligned} \tag{A.19}$$

The same process gives

$$
\begin{aligned}
S2 &= \sum_{l=0}^{N}\sum_{k=0}^{N} f(k)g(\ell)\left[\sum_{s=1}^{2N}\cos[k(\phi_v-\phi_s)]\sin(\ell\phi_s)\right] \\
&= \sum_{l=0}^{N}\sum_{k=0}^{N} f(k)g(\ell)\ N(\delta_{k-\ell}-\delta_{k+\ell}-\delta_{k+\ell-2N})\sin(k\phi_v) \\
&= \sum_{l=0}^{N} f(\ell)g(\ell)\ N(1-\delta_{2\ell}-\delta_{2\ell-2N})\sin(\ell\phi_v) \\
&= \sum_{l=0}^{N} \gamma_\ell f(\ell)g(\ell)\sin(\ell\phi_v)
\end{aligned}
\tag{A.20}
$$

Likewise

$$
\begin{aligned}
S3 &= \sum_{l=0}^{N}\sum_{k=0}^{N} f(k)g(\ell)\left[\sum_{s=1}^{2N}\sin[k(\phi_v-\phi_s)]\cos(\ell\phi_s)\right] \\
&= \sum_{l=0}^{N} \epsilon_\ell f(\ell)g(\ell)\sin(\ell\phi_v)
\end{aligned}
\tag{A.21}
$$

and

$$
\begin{aligned}
S4 &= \sum_{l=0}^{N}\sum_{k=0}^{N} f(k)g(\ell)\left[\sum_{s=1}^{2N}\sin[k(\phi_v-\phi_s)]\sin(\ell\phi_s)\right] \\
&= -\sum_{l=0}^{N} \gamma_\ell f(\ell)g(\ell)\cos(\ell\phi_v)
\end{aligned}
\tag{A.22}
$$

Note the minus sign in the last equation.

These results for the Fourier decomposition of discrete functions of one or two azimuthal angles provide all of the tools necessary for converting the discretized physical-space VRTE into discretized equations in Fourier space.

APPENDIX B

Ray Tracing in a Hexagonal Sea Surface Grid

This appendix gives the detailed algorithm for tracing rays through 3D space as they interact with random sea surfaces.

Section §3.2.1 showed how to model a wind-blown sea surface as a grid of triangular wave facets, each of which is locally plane but tilted with respect to the normal to the mean sea surface. Once such a sea-surface grid is generated, we must be able to follow a photon ray path as the ray first approaches the sea surface from above or below, intersects a sea-surface wave facet, and is then reflected and refracted to generate daughter rays. A ray incident onto the surface from the air side always generates two daughter rays, one reflected back to the air and one refracted through the surface into the water. A ray incident from the water side may generate both reflected and refracted rays, or it may generate only a reflected ray back into the water if the angle of incidence onto the underside of the surface facet is greater than the critical angle for total internal reflection. As explained in §3.2.2, these daughter rays are pushed into a stack array as they are generated, and the rays in the stack are processed until the stack is empty. The daughter rays that leave the hexagonal domain are tallied and contribute to the estimate of a surface reflectance or transmission function.

The points of ray-surface intersections and the ray paths through 3D space depend on the exact shape of the random sea surface and on the incident ray direction. The mathematics of following ray paths through 3D space in the region of the sea surface is in principle no more complicated than determining where a straight line intersects a plane. However, the application of the basic geometric concepts to the irregular geometry of the wave facets comprising a random realization of the sea surface is quite tedious and requires careful attention to detail. A ray-tracing algorithm should should be applicable to any sea surface shape and should allow for multiple ray interactions with the surface. Such an algorithm was first worked out by Preisendorfer and Mobley (1985) and has been in use for almost three decades for the surface calculations in the HydroLight software (which uses Cox-Munk sea surfaces, generated as described in §3.2.1). The same algorithm is adopted in HydroPol. However, the algorithm was documented in a technical report that is long-since out of print. It is therefore necessary for completeness of the HydroPol documentation to describe again the ray-tracing algorithm. The figures in this appendix are taken (with minor modification) from the Preisendorfer and Mobley (1985) report.

The directions of an incident ray and the normal to an intersected ray facet determine a scattering plane that contains the incident, reflected, and refracted directions and the normal to the facet, as is seen in Fig. 3.13. Those calculations are purely geometric and do not depend on polarization. However, when the ray paths are used to compute the reflection and refraction of polarized light by the sea surface, the rotations of Stokes vectors from one facet scattering plane to the next must be included. How that is done, given the ray paths as determined here, is explained in §3.2.2. This appendix considers only the geometric calculations needed to trace rays through 3D space as they interact with a random sea surface.

The coordinate system is the same as in Fig. 1.1. The origin of the **x**-**y**-**z** system is placed at mean sea level in the center of the hexagonal grid of surface triads as in Fig. B.1. Let $\boldsymbol{\xi}'$ be the direction of a light ray incident onto the sea surface from above or below. This parent ray $\boldsymbol{\xi}'$ has components as shown in Eq. (1.3). Likewise, $\mathbf{n}$ is the outward normal to a wave facet, as seen in Fig. 3.13 (outward normal means the orientation of $\mathbf{n}$ for which $\mathbf{n} \cdot \mathbf{z} > 0$).

Figure B.1 illustrates a hexagonal grid with $N_{\text{hex}} = 2$ concentric rings or layers of triads encircling the origin. As seen in the inset at the lower right, each triad has a base length of δ in the $\mathbf{x}$ (alongwind) direction and a height of ϵ in the $\mathbf{y}$ (crosswind) direction. Triad side $\gamma = [(\delta/2)^2 + \epsilon^2]^{1/2}$ is also defined in the inset. These dimensions are determined by the spatial dimensions of the patch of sea surface being simulated and the number of grid points covering that area. The lines forming the sides of the triads have unit vector directions given by

$$\begin{aligned}
\mathbf{r}_0 &= (1, 0, 0) = \mathbf{x} \\
\mathbf{r}_1 &= (r_{11}, r_{21}, 0) = \frac{\delta}{2\gamma}\mathbf{x} + \frac{\epsilon}{\gamma}\mathbf{y} \\
\mathbf{r}_2 &= (r_{12}, r_{22}, 0) = -\frac{\delta}{2\gamma}\mathbf{x} + \frac{\epsilon}{\gamma}\mathbf{y}
\end{aligned}$$

Note that

$$\tilde{\mathbf{r}}_k = (-r_{2k}, r_{1k}, 0) \tag{B.1}$$

is a vector orthogonal to $\mathbf{r}_k$ for $k = 1$ or 2.

A hexagonal grid with N_{hex} layers of triads has $3N_{\text{hex}}(N_{\text{hex}} + 1)$ triad vertices or nodes where the elevations of the facet are defined. These nodes are located at $\pm b\delta\mathbf{x} + c\epsilon\mathbf{y}$ when c is even, for $b = 0, 1, ..., N_{\text{hex}} - c/2$; and at $\pm(b + 1/2)\delta\mathbf{x} + c\epsilon\mathbf{y}$ when c is odd, for $b = 0, 1, ..., N_{\text{hex}} - (c + 1)/2$, where in either the even or odd case, $c = 0, \pm 1, \pm 2, ..., \pm N_{\text{hex}}$.

The lines forming the sides of the triads are specified in terms of $\mathbf{r}_0, \mathbf{r}_1$, and $\mathbf{r}_2$ as follows. The "r_0" family of horizontal lines in Fig. B.1 is a family of lines, each of which is a set of points of the form

$$\lambda_0\mathbf{x} + c\epsilon\mathbf{y}$$

where λ_0 is a real number. That is, the line is generated as λ_0 ranges over the set of real numbers for some fixed value of c. For example, $c = 0$ gives for the horizontal line through the origin, $c = N_{\text{hex}}$ gives the line forming the top boundary of the hexagon, and $c = -N_{\text{hex}}$ gives the line along the lower boundary of the hexagon. The members of the "r_1" and "r_2" families of lines are given in the same fashion by

$$\begin{aligned}
&\lambda_1\mathbf{r}_1 + 2c\epsilon\mathbf{y} \\
&\lambda_2\mathbf{r}_2 + 2c\epsilon\mathbf{y}\,.
\end{aligned}$$

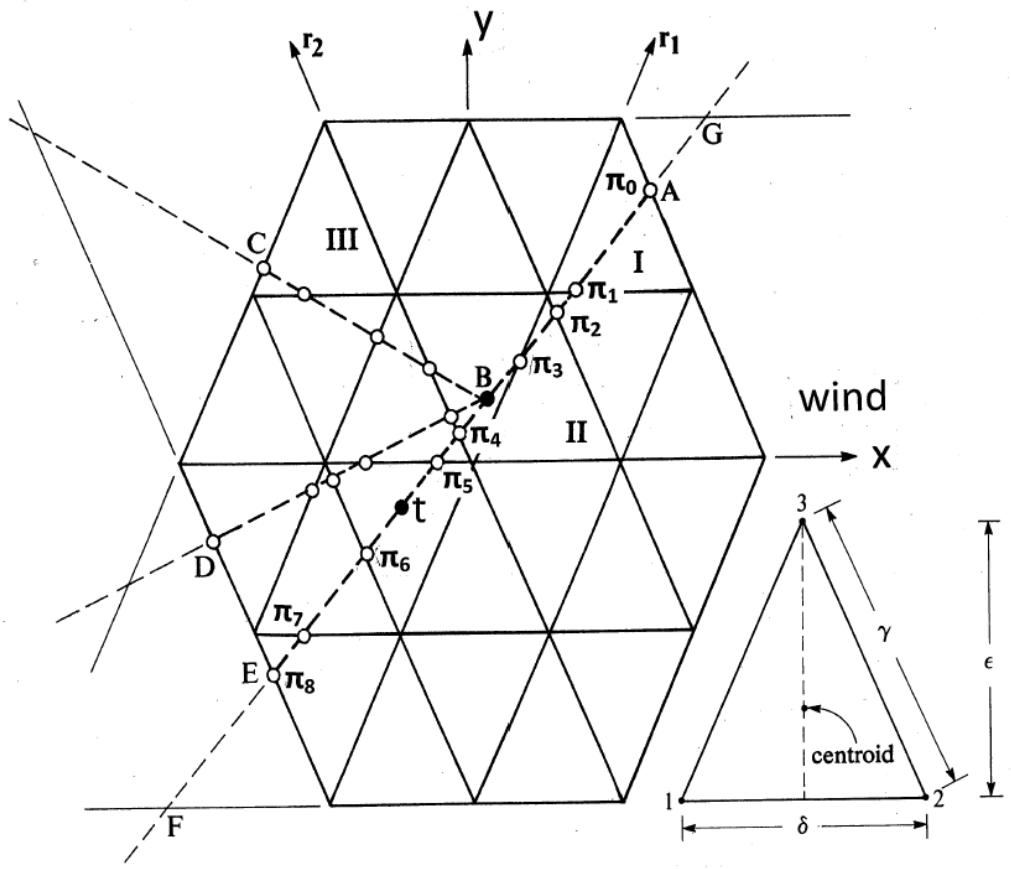

Figure B.1: The wind-based hexagonal grid for $N_{\text{hex}} = 2$. Example ray tracks are shows as dashed lines AE, BC, BD. The triad intercept points (TIPs) are show by open circles; the TIPs along track AE are labeled $\boldsymbol{\pi}_0$ to $\boldsymbol{\pi}_8$. The inset at the lower right shows the triad dimensions.

for λ_1 and λ_2 real and some fixed value of c as before.

A ray *path* is the set of points in 3D space taken by a photon as it travels through space. A ray path is specified by a ray direction unit vector $\boldsymbol{\xi} = (\sin\theta\cos\phi, \sin\theta\sin\phi, \cos\theta)$. A ray *track* is the projection of a ray path onto the 2D hexagonal grid. Thus a ray track is a horizontal ($\theta = 90$ deg) line segment lying in the hexagonal grid and is given by $\boldsymbol{\xi}_h = (\cos\phi, \sin\phi, 0)$. The dashed lines in Fig. B.1 show several ray tracks. Track AE represents a ray whose projection enters the hexagonal grid a boundary point A. The ray is traveling in space toward a target point in the grid at $\mathbf{t} = (-\delta/2, -\epsilon/3, 0)$. If this ray never intersected the surface it would leave the grid at point E. Suppose the ray intersects the wave surface at a point below or above point B before reaching the target point $\mathbf{t}$. The ray then generates reflected and refracted daughter rays that have tracks BC and BD.

The points where a track crosses the sides of the triads are called *triad intercept points* or TIPs. For trace AE these points are labeled $\boldsymbol{\pi}_0, \boldsymbol{\pi}_1, ..., \boldsymbol{\pi}_8$. If a ray track crosses a triad, there are usually two triad intercept points, and we can characterize the crossing by specifying which sides of the triad have the TIPs. Thus triad I in B.1 is said to be of type "$r_0 r_2$" because on triad intercept

point, $\boldsymbol{\pi}_1$, lies on a side that is a member of the r_0 family of lines, and the other TIP, $\boldsymbol{\pi}_0$, lies on a side that is a member of the r_2 family of lines. Triad II is said to be of type "$r_1 r_2$," and triad III is an example of type "$r_0 r_1$".

In analyzing the interactions between a ray and the wave facets, two problems must be solved. First, in the case of an incoming ray, we must determine where the ray first enters the hexagonal domain (e.g., point $\mathbf{p}'$ in Fig. B.2, corresponding to point A in Fig. 3.6). In the case of a daughter ray, we must determine where the ray leaves the hexagonal domain (e.g., points D, E, and G in Fig. 3.6 or points C and D in Fig. B.1). Second, we must determine where, if anywhere, the ray intercepts a wave facet (e.g., points B, C, and F of Fig. 3.6). The next two sections give the solutions to these two problems.

B.1 Determination of the Initial Ray Point

A ray is defined by an initial point $\mathbf{p} = (p_x, p_y, p_z)$ and a direction $\boldsymbol{\xi}$. An initial point can be either on the boundary of the hexagonal domain (the case of an incoming ray), or in the interior of the domain (the case of a reflected or refracted daughter ray emanating from the water surface). When initializing a new ray with direction $\boldsymbol{\xi}'$ in the Monte Carlo calculations, it is convenient to select a target point $\mathbf{t}$ near the center of the hexagonal grid, toward which the ray $\boldsymbol{\xi}'$ is directed. Therefore, for a given initial direction $\boldsymbol{\xi}'$ and target point $\mathbf{t}$, we must first locate the point $\mathbf{p}'$ where the incoming ray $\boldsymbol{\xi}'$ enters the hexagonal domain on its way to $\mathbf{t}$. The ray $\boldsymbol{\xi}'$ can then be traced in the normal fashion starting at initial point $\mathbf{p}'$ until it intersects the surface somewhere on its way to $\mathbf{t}$.

Consider now an incoming ray at $\mathbf{p}'$ traveling in direction $\boldsymbol{\xi}'$ toward target point $\mathbf{t}$, as shown in Fig. B.2. The track of ray $\boldsymbol{\xi}'$ is the straight line containing points of the form $\mathbf{t} + s\boldsymbol{\xi}'_h$, where $\boldsymbol{\xi}'_h = (\cos\phi', \sin\phi', 0)$ is the projection of $\boldsymbol{\xi}'$ onto the hexagonal grid, and $\mathbf{p}'_h = (p'_x, p'_y, 0)$, the projection of $\mathbf{p}'$ onto the grid. s is a real number representing the distance traveled in the grid away from the target point $\mathbf{t}$. As s varies, the point moves along the track. By choice, s is measured as positive from 0 at $\mathbf{t}$ when moving along the track in the $-\boldsymbol{\xi}'$ direction. This track is illustrated as the infinite extension of the dashed line segment AE in Figs. B.1 and B.2.

Note in Fig. B.1 that the track first enters the hexagonal grid at A, where the track crosses a boundary line of the r_2 family. The track has previously crossed an r_0 family boundary line at point G, and the r_1 family boundary line was first crossed at a point outside the figure. The track also crosses an r_0 family boundary line at point F, an r_2 family boundary line at E, and an r_1 family boundary line at a point not shown.

In general a track $\mathbf{t} + s\boldsymbol{\xi}_h$ intersects the r_0 family of lines at a set of points whose distances s from $\mathbf{t}$ can be found by setting

$$\mathbf{t} + s\boldsymbol{\xi}_h = \lambda_0 \mathbf{x} + c\epsilon \mathbf{y}$$

and solving for s. Taking the dot product with $\mathbf{y}$ gives

$$\mathbf{t} \cdot \mathbf{y} + s\boldsymbol{\xi}_h \cdot \mathbf{y} = c\epsilon\,,$$

so that the s values are given by

$$s_0(c, \mathbf{t}, \boldsymbol{\xi}_h) = \frac{c\epsilon - \mathbf{t} \cdot \mathbf{y}}{\boldsymbol{\xi}_h \cdot \mathbf{y}}\,, \tag{B.2}$$

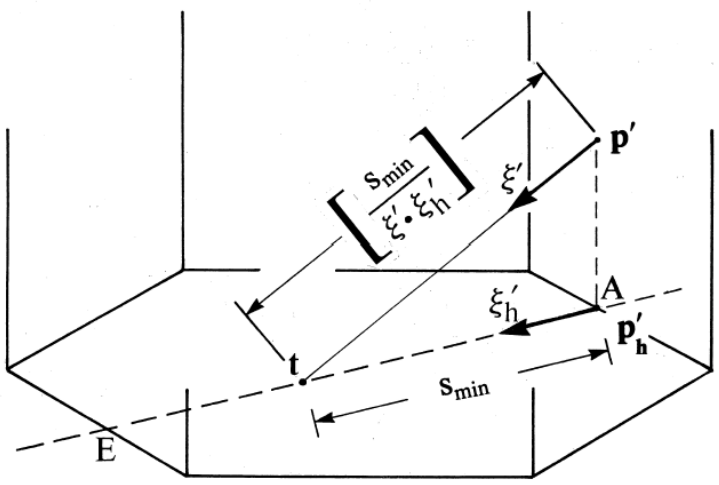

Figure B.2: Geometry for determination of the initial point $\mathbf{p}'$ at which an incident ray $\boldsymbol{\xi}'$ enters the hexagonal domain. $\boldsymbol{\xi}'_h$ is the projection of $\boldsymbol{\xi}'$ onto the hexagonal grid of triads; likewise, $\mathbf{p}' = (p'_x, p'_y, 0)$. The hexagonal domain is the cylindrical region in 3D space formed by a vertical line tracing the perimeter of the hexagonal grid.

with $c = 0, \pm 1, \pm 2, ..., \pm N_{\text{hex}}$ and $\boldsymbol{\xi}_h \cdot \mathbf{y} \neq 0$. Likewise the intersections of the track with the r_1 and r_2 families of lines are found by setting

$$\mathbf{t} + s\boldsymbol{\xi}_h = \lambda_k \mathbf{r}_k + 2c\epsilon\mathbf{y}, \quad k = 1, 2$$

and taking the dot product with the $\tilde{\mathbf{r}}_k$ of Eq. (B.1) to get

$$s_k(c, \mathbf{t}, \boldsymbol{\xi}_h) = \frac{2c\epsilon\mathbf{y} \cdot \tilde{\mathbf{r}}_k - \mathbf{t} \cdot \tilde{\mathbf{r}}_k}{\boldsymbol{\xi}_h \cdot \tilde{\mathbf{r}}_k}\,, \tag{B.3}$$

for $k = 1, 2; c = 0, \pm 1, \pm 2, ..., \pm N_{\text{hex}}$ and $\xi_h \cdot \tilde{\mathbf{r}}_k \neq 0$. These s_k are the distances from $\mathbf{r}$ at which the track crosses the r_k families of lines. As above, s_k is measured positive from $\mathbf{t}$ along $-\boldsymbol{\xi}_h$.

If a ray is started from $\mathbf{t}$ in a direction $-\boldsymbol{\xi}_h$, then its track crosses the boundary lines of the r_k family at $s_k(\pm N_{\text{hex}}, \mathbf{t}, -\boldsymbol{\xi}_h)$. In the example of Fig. B.1, such a track proceeds from $\mathbf{t}$ towards A (i.e., in the $-\boldsymbol{\xi}_h$ direction), building up positive s values until the r_2 boundary line is crossed at point A a distance $s_2(N_{\text{hex}}, \mathbf{t}, -\boldsymbol{\xi}_h)$ from $\mathbf{t}$. The r_0 boundary line is then crossed at point G a greater distance $s_0(N_{\text{hex}}, \mathbf{t}, -\boldsymbol{\xi}_h)$, and finally the r_1 boundary line is crossed at some still greater distance $s_1(N_{\text{hex}}, \mathbf{t}, -\boldsymbol{\xi}_h)$ from $\mathbf{t}$. In this example, we decide that the distance from $\mathbf{t}$ to the hexagonal grid boundary is the positive number $s_2(N_{\text{hex}}, \mathbf{t}, -\boldsymbol{\xi}_h)$. We do not choose points at E and F of Fig. B.1 to define the distance because they have negative values $s_2(-N_{\text{hex}}, \mathbf{t}, -\boldsymbol{\xi}_h)$ and $s_0(-N_{\text{hex}}, \mathbf{t}, -\boldsymbol{\xi}_h)$, respectively, measured from $\mathbf{t}$. In general, the six values $s_k(\pm N_{\text{hex}}, \mathbf{t}, -\boldsymbol{\xi}_h), k = 0, 1, 2$, define points at which a track from $\mathbf{t}$ crosses the hexagonal grid boundary lines. Following the idea in the example just cited, we define the distance from $\mathbf{t}$ to the boundary of the hexagonal grid as

$$s_{\min}(\mathbf{t}, -\boldsymbol{\xi}_h) = \text{minimum of the positive values of } s_k(\pm N_{\text{hex}}, \mathbf{t}, -\boldsymbol{\xi}_h), \quad k = 0, 1, 2\,.$$

Then the initial point $\mathbf{p}'$ at which the ray $\boldsymbol{\xi}'$ enters the hexagonal domain on its way to $\mathbf{t}$ is given

by

$$\mathbf{p}' = \mathbf{t} - \left[\frac{s_{\min}(\mathbf{t}, -\boldsymbol{\xi}_h)}{\boldsymbol{\xi}' \cdot \boldsymbol{\xi}_h}\right] \boldsymbol{\xi}', \quad \boldsymbol{\xi}' \cdot \boldsymbol{\xi}_h \neq 0 \tag{B.4}$$

as can be seen geometrically from Fig. B.2.

Using the initial point $\mathbf{p}'$ from Eq. (B.4), we can now proceed with the ray tracing calculations for initial ray $\boldsymbol{\xi}'$.

B.2 Determination of Triad Intersection Points

Let $\mathbf{p} = (p_x, p_y, p_z)$ be any point in the hexagonal grid or on its boundary, and let $\boldsymbol{\xi}$ be an arbitrary direction at $\mathbf{p}$. In the following development, it is assumed that $\mathbf{p}$ and $\boldsymbol{\xi}$ are held fixed. The projection of $\mathbf{p}$ onto the hexagonal grid is $\mathbf{p}_h = (p_x, p_y, 0)$, and $\boldsymbol{\xi}_h$ is the unit vector determined as before from the projection of $\boldsymbol{\xi}$. Then $\mathbf{p}_h + s\boldsymbol{\xi}_h$, as s varies from $\mathbf{p}_h$ to the intercepts of the track with the families of the lines defining the triads are then given by equations analogous to (B.2) and (B.3):

$$s_0(c, \mathbf{p}_h, \boldsymbol{\xi}_h) = \frac{c\epsilon - \mathbf{p}_h \cdot \mathbf{y}}{\boldsymbol{\xi}_h \cdot \mathbf{y}} \tag{B.5a}$$

$$s_k(c, \mathbf{p}_h, \boldsymbol{\xi}_h) = \frac{2c\epsilon \mathbf{y} \cdot \tilde{\mathbf{r}}_k - \mathbf{p}_h \cdot \tilde{\mathbf{r}}_k}{\boldsymbol{\xi}_h \cdot \tilde{\mathbf{r}}_k}, \quad k = 1, 2 \tag{B.5b}$$

where $c = 0, \pm 1, \pm 2, ..., \pm N_{\text{hex}}$. Following the pattern of definition of $s_{\min}(\mathbf{t}, -\boldsymbol{\xi}_h)$, the distance from $\mathbf{p}_h$ to the boundary of the hexagon is now defined as

$$s_{\min}(\mathbf{p}_h, \boldsymbol{\xi}_h) = \text{minimum of the positive values of } s_k(\pm N_{\text{hex}}, \mathbf{p}_h, \boldsymbol{\xi}_h), \quad k = 0, 1, 2\,. \tag{B.6}$$

Equations (B.5) define the distances (positive along $\boldsymbol{\xi}_h$) from $\mathbf{p}_h$ to all possible intersections of the track with the lines that define the triad boundaries. However, only those $s_k(c, \mathbf{p}_h, \boldsymbol{\xi}_h)$ values between 0 and $s_{\min}(\mathbf{p}_h, \boldsymbol{\xi}_h)$ are within the hexagon. For track AE of Fig. B.1, the triad intercept point (TIP) $\boldsymbol{\pi}_0$ is at a distance $s_2(2, \mathbf{p}_h, \boldsymbol{\xi}_h) = 0$ from A, and so on until $\boldsymbol{\pi}_8$ at E is reached at a distance $s_{\min}(\mathbf{p}_h, \boldsymbol{\xi}_h) = s_2(-2, \mathbf{p}_h, \boldsymbol{\xi}_h)$ from A. Note that as we proceed from $\mathbf{p}_h$ to the boundary of the hexagon, the $s_k(c, \mathbf{p}_h, \boldsymbol{\xi}_h)$ values can occur in any order of $k = 0, 1$ or 2 in the encountered sequence of triad intercept points $\boldsymbol{\pi}_j, j = 0, 1, ..., m$. Thus in Fig. B.1, $\boldsymbol{\pi}_0$ is associated with an s value from the $k = 2$ family, $\boldsymbol{\pi}_1$ with $k = 0$, $\boldsymbol{\pi}_2$ with $k = 2$, $\boldsymbol{\pi}_3$ with $k = 1$, and so on. It will be necessary to know which k family of distances $s_k(c, \mathbf{p}_h, \boldsymbol{\xi}_h)$ is associated with each TIP $\boldsymbol{\pi}_j$. Toward this end, let $s(j)$ denote the $s_k(c, \mathbf{p}_h, \boldsymbol{\xi}_h)$ value associated with TIP $\boldsymbol{\pi}_j$, and let $\kappa(j)$ denote the k family of $s(j)$. Thus $\kappa(j) = 0, 1$, or 2, depending on whether $k = 0, 1$, or 2 for the $s_k(c, \mathbf{p}_h, \boldsymbol{\xi}_h)$ value belonging to $\boldsymbol{\pi}_j$. With this notation, the triad intercept points are given by

$$\boldsymbol{\pi}_j = \mathbf{p}_h + s(j)\boldsymbol{\xi}_h, \quad j = 0, 1, ..., m\,. \tag{B.7}$$

The maximum possible value of m is $4N_{\text{hex}}$, an example of which is track AE in Fig. B.1 (track BC has $m = 3$ and track BD has $m = 4$).

In practice the TIPs are found by the following steps:

1. Generate all possible $s_k(c, \mathbf{p}_h, \boldsymbol{\xi}_h)$ values by Eq. (B.5).

2. Find $s_{\min}(\mathbf{p}_h, \boldsymbol{\xi}_h)$ by Eq. (B.6).

3. Record all s_k distances, as c varies, such that $0 \le s_k(c, \mathbf{p}_h, \boldsymbol{\xi}_h) \le s_{\min}(\mathbf{p}_h, \boldsymbol{\xi}_h)$.

4. Order the values of s_k from Step (3) from smallest to largest.

5. Relabel these ordered s_k values as $s(j)$, with $j = 0$ for the smallest value from Step (4), namely $s(0)$ for the smallest $s_k(c, \mathbf{p}_h, \boldsymbol{\xi}_h)$, to $s(m) = s_{\min}(\mathbf{p}_h, \boldsymbol{\xi}_h)$ for the largest value.

6. Record the $\kappa(j)$ value associated with each j value of Step (5).

7. Generate the ordered TIP values $\boldsymbol{\pi}_j$ using Eq. (B.7).

B.3 Determination of the Triad Vertices

We now show how to recover the locations of the triad vertices from the triad intercept points just computed. This capability enables us to identify those facets that are candidates for interaction with the ray, since triads (and their associated wave facets) are uniquely determined by the coordinates of the vertices.

Consider a triad that is crossed by a track of type $r_1 r_2$. Triad II of Fig. B.1 is an example. Figure B.3 illustrates in greater detail the geometry of an $r_1 r_2$ triad. The triad intercept points $\boldsymbol{\pi}_{j_1} = \boldsymbol{\pi}_1$ and $\boldsymbol{\pi}_{j_2} = \boldsymbol{\pi}_2$ locate an arbitrary pair of type $r_1 r_2$ TIPs (i.e., either $j_2 = j_1 + 1$ or $j_1 = j_2 + 1$). The notation $\boldsymbol{\pi}_1, \boldsymbol{\pi}_2$ is used only for convenience to label which TIP $\boldsymbol{\pi}_k$ lies on which side $\mathbf{r}_k$ of the triad. The TIPs $\boldsymbol{\pi}_1$ and $\boldsymbol{\pi}_2$ have the associated distances $s(1) = s(j_1)$ and $s(2) = s(j_2)$, respectively, (as in Eq. (B.7). Clearly, $\kappa(j_1) + \kappa(j_2) = \kappa(1) + \kappa(2) = 3$ if and only if the triad crossing is of type $r_1 r_2$.

Since TIP $\boldsymbol{\pi}_k$ lies on a line of the $\mathbf{r}_k$ family, $k = 1$ or 2, we can write, with the help of Fig. B.3:

$$\boldsymbol{\pi}_k = \mathbf{p}_h + s(k)\boldsymbol{\xi}_h = \lambda_k \mathbf{r}_k + d_k \mathbf{y} \quad (r_1 r_2 \text{ case}) .$$

The rightmost of these equations is illustrated in Fig. B.3 for $k = 1$. d_1 and d_2 are defined in the figure. Taking the dot product with the $\tilde{\mathbf{r}}_k$ of Eq. (B.1) yields d_k:

$$d_k = \frac{[\mathbf{p}_h + s(k)\boldsymbol{\xi}_h] \cdot \tilde{\mathbf{r}}_k}{\mathbf{y} \cdot \tilde{\mathbf{r}}_k} = \frac{\boldsymbol{\pi}_k \cdot \tilde{\mathbf{r}}_k}{\mathbf{y} \cdot \tilde{\mathbf{r}}_k}, \quad k = 1, 2 \quad (r_1 r_2 \text{ case}) . \tag{B.8}$$

Since vertex $\mathbf{a}$, as shown in Fig. B.3, lies on the $\mathbf{r}_k$ lines, we have for some $\lambda_k = \lambda_k(\mathbf{a})$,

$$\lambda_k(\mathbf{a})\mathbf{r}_k + d_k \mathbf{y} = \mathbf{a} , \quad k = 1, 2 \quad (r_1 r_2 \text{ case})$$

From this we find

$$\lambda_k(\mathbf{a}) = [\mathbf{a} - d_k \mathbf{y}] \cdot \mathbf{r}_k , \quad k = 1, 2 \quad (r_1 r_2 \text{ case}) \tag{B.9}$$

Since $\boldsymbol{\pi}_k$ lies on an $\mathbf{r}_k$ line, we find in the same way that

$$\lambda_k = [\boldsymbol{\pi}_k - d_k \mathbf{y}] \cdot \mathbf{r}_k , \quad k = 1, 2 \quad (r_1 r_2 \text{ case}) \tag{B.10}$$

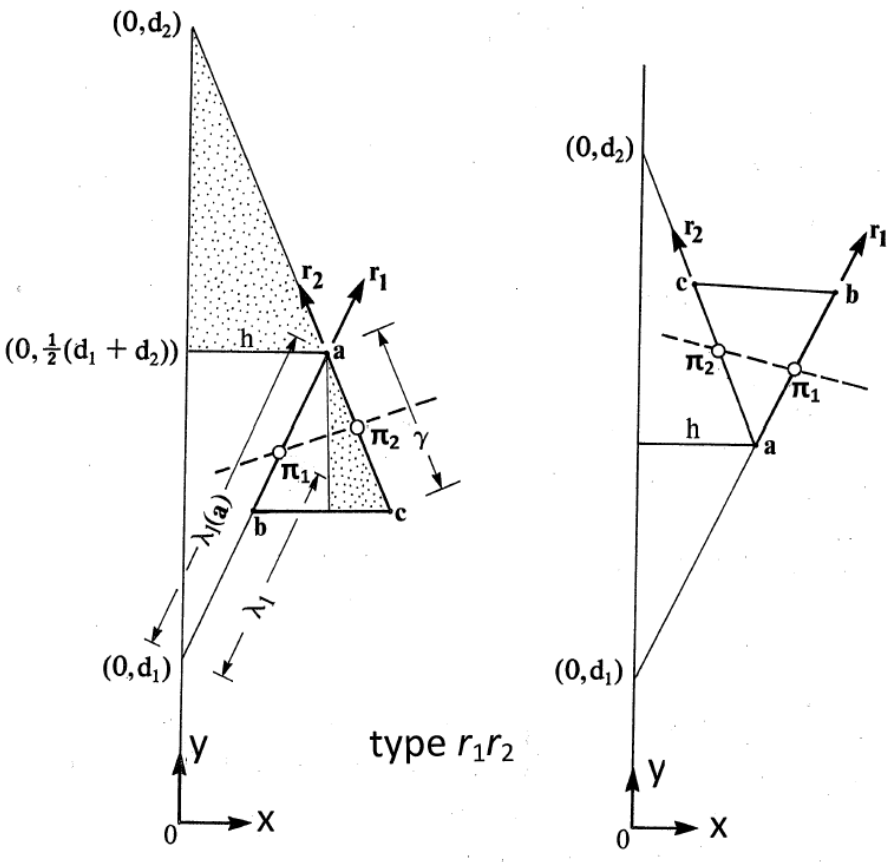

Figure B.3: Geometric relations for determination of the triad vertices **a**, **b**, and **c** from triad intercept points $\boldsymbol{\pi}_1$ and $\boldsymbol{\pi}_2$ that are of type $r_1 r_2$.

Note that d_k, λ_k, and $\lambda_k(\mathbf{a})$ can be positive or negative, depending on the location of the triad in the hexagonal grid. The quantities d_k, for $k = 1, 2$, along with λ_1 and $\lambda_1(\mathbf{a})$ are explicitly shown in Fig. B.3. From similar triangles, shown shaded in Fig. B.3, we have

$$\frac{h}{\frac{1}{2}(d_2 - d_1)} = \frac{\frac{1}{2}\delta}{\epsilon}$$

or

$$h = \frac{(d_2 - d_1)\delta}{4\epsilon}\,, \tag{B.11}$$

where h is defined geometrically in Fig. B.3, and the triad dimensions δ, ϵ are shown in Fig. B.1.

We can now write **a** in terms of quantities computed from the TIPs:

$$\mathbf{a} = h\mathbf{x} + \frac{1}{2}(d_1 + d_2)\mathbf{y}\,.$$

Defining

$$\text{sign[x]} \equiv \begin{cases} +1 & \text{if } x > 0 \\ -1 & \text{if } x < 0 \end{cases}$$

we can write **b** and **c** as

$$\mathbf{b} = \mathbf{a} + \text{sign}[\lambda_1 - \lambda_1(\mathbf{a})]\gamma\mathbf{r}_1$$
$$\mathbf{c} = \mathbf{a} + \text{sign}[\lambda_2 - \lambda_2(\mathbf{a})]\gamma\mathbf{r}_2$$

where γ is shown in Figs. B.1 and B.3. Recalling the definitions of

$$\mathbf{r}_1 = \frac{\delta}{2\gamma}\mathbf{x} + \frac{\epsilon}{\gamma}\mathbf{y}$$
$$\mathbf{r}_2 = -\frac{\delta}{2\gamma}\mathbf{x} + \frac{\epsilon}{\gamma}\mathbf{y}$$

we can write points **a**, **b**, and **c** in terms of the $\mathbf{x}, \mathbf{y}$ axis unit vectors and the quantities computed from the TIPs:

$$\mathbf{a} = h\mathbf{x} + \frac{1}{2}(d_1 + d_2)\mathbf{y} \equiv (a_1, a_2, 0) \quad (r_1r_2 \text{ case}) \tag{B.12a}$$

$$\mathbf{b} = \left(a_1 + \text{sign}[\lambda_1 - \lambda_1(\mathbf{a})]\frac{\delta}{2}\right)\mathbf{x} + (a_2 + \text{sign}[\lambda_1 - \lambda_1(\mathbf{a})]\epsilon)\,\mathbf{y} \tag{B.12b}$$

$$\mathbf{c} = \left(a_1 - \text{sign}[\lambda_2 - \lambda_2(\mathbf{a})]\frac{\delta}{2}\right)\mathbf{x} + (a_2 + \text{sign}[\lambda_2 - \lambda_2(\mathbf{a})]\epsilon)\,\mathbf{y} \tag{B.12c}$$

Observe that these equations hold for both possible triad orientations of the r_1r_2 case shown in Fig. B.3, as well as for the instances where the triad is to the left of the origin (in which case h is negative). Therefore, given a pair of TIPs $\boldsymbol{\pi}_{j_1}$ and $\boldsymbol{\pi}_{j_2}$, we need only check the sum $\kappa(j_1) + \kappa(j_2)$. If this sum is 3, then Eqs. (B.8) to (B.12) give the vertices of the triad.

Figure B.4 shows the geometries of the remaining two types of triad intercepts, namely r_0r_1 and r_0r_2. Labeling the TIP that is on the r_0 side of the triad as $\boldsymbol{\pi}_0$, with $\boldsymbol{\pi}_k$ denoting as above the TIP on the $\mathbf{r}_k$ side ($k = 1$ or 2), the analysis proceeds just in the r_1r_2 case. The relevant equations are as follows.

For triad intercept type r_0r_1, for which $\kappa(j_0) + \kappa(j_1) = 1$, we have (since $\boldsymbol{\pi}_0$ lies on an **x** line)

$$\boldsymbol{\pi}_0 = \mathbf{p}_h + s(0)\boldsymbol{\xi}_h = \lambda_0\mathbf{x} + d_0\mathbf{y}\,,$$

whence

$$d_0 = \boldsymbol{\pi}_0 \cdot \mathbf{y}\,.$$

Moreover, since **a**, as shown in Fig. 63-8, also lies on an **x** line, we can write

$$\mathbf{a} = \lambda_0(\mathbf{a})\mathbf{x} + d_0\mathbf{y}\,,$$

whence

$$\lambda_0(\mathbf{a}) = \mathbf{a} \cdot \mathbf{x}\,.$$

Similarly,

$$\lambda_0 = \boldsymbol{\pi}_0 \cdot \mathbf{x}\,.$$

Since $\boldsymbol{\pi}_1$ lies on an r_1 line, we have d_1 given by Eq. (B.8), $\lambda_1(\mathbf{a})$ by (B.9), and λ_1 by (B.10). As before, similar triangles give

$$\frac{h}{d_0 - d_1} = \frac{\frac{1}{2}\delta}{\epsilon}$$

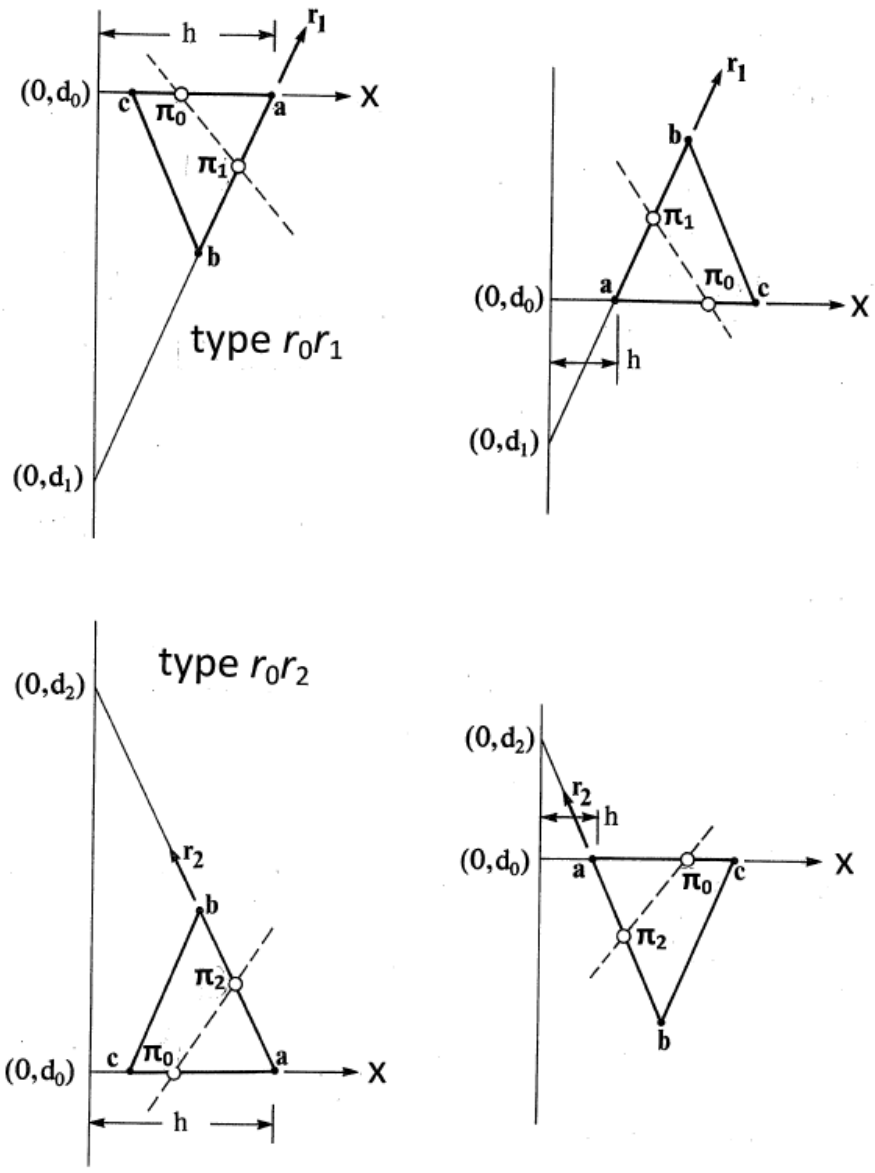

Figure B.4: Geometric relations as in Fig. B.3 for triad intercept types r_0r_1 (top) and r_0r_2 (bottom).

or

$$h = \frac{(d_0 - d_1)\delta}{2\epsilon}$$

The final equations for the triad vertices of type r_0r_1 or $\kappa(j_0)+\kappa(j_1)=1$ are then

$$\mathbf{a} = h\mathbf{x} + d_0\mathbf{y} \equiv (a_1, a_2, 0) \quad (r_0r_1 \text{ case}) \tag{B.13a}$$

$$\begin{aligned}\mathbf{b} &= \mathbf{a} + \text{sign}[\lambda_1 - \lambda_1(\mathbf{a})]\gamma\mathbf{r}_1 \\ &= \left(a_1 + \text{sign}[\lambda_1 - \lambda_1(\mathbf{a})]\frac{\delta}{2}\right)\mathbf{x} + (a_2 + \text{sign}[\lambda_1 - \lambda_1(\mathbf{a})]\epsilon)\,\mathbf{y}\end{aligned} \tag{B.13b}$$

$$\begin{aligned}\mathbf{c} &= \mathbf{a} + \text{sign}[\lambda_0 - \lambda_0(\mathbf{a})]\delta\mathbf{x} \\ &= (a_1 + \text{sign}[\lambda_0 - \lambda_0(\mathbf{a})]\delta)\,\mathbf{x} + a_2\mathbf{y}\,.\end{aligned} \tag{B.13c}$$

For triad intercept type r_0r_2, for which for which $\kappa(j_0)+\kappa(j_2)=2$, an identical analysis leads to the final equations

$$\mathbf{a} = h\mathbf{x} + d_0\mathbf{y} \equiv (a_1, a_2, 0) \quad (r_0r_2 \text{ case}) \tag{B.14a}$$

$$\begin{aligned}\mathbf{b} &= \mathbf{a} + \text{sign}[\lambda_2 - \lambda_2(\mathbf{a})]\gamma\mathbf{r}_2 \\ &= \left(a_1 - \text{sign}[\lambda_2 - \lambda_2(\mathbf{a})]\frac{\delta}{2}\right)\mathbf{x} + (a_2 + \text{sign}[\lambda_2 - \lambda_2(\mathbf{a})]\epsilon)\,\mathbf{y}\end{aligned} \tag{B.14b}$$

$$\begin{aligned}\mathbf{c} &= \mathbf{a} + \text{sign}[\lambda_0 - \lambda_0(\mathbf{a})]\delta\mathbf{x} \\ &= (a_1 + \text{sign}[\lambda_0 - \lambda_0(\mathbf{a})]\delta)\,\mathbf{x} + a_2\mathbf{y}\,.\end{aligned} \tag{B.14c}$$

Here, as before, d_2 is given by Eq. (B.8), $\lambda_2(\mathbf{a})$ by (B.9), and λ_2 by (B.10).

Given any possible geometry for a track crossing a triad, we can now immediately obtain the vertices of the triad by either Eqs. (B.12), (B.13), or (B.14).

B.4 Determination of the Facet Intersection Point

The elevations of the random sea surface are defined at the nodes of the hexagonal grid, i.e., at the triad vertices. Given the triad vertices $\mathbf{a}$, $\mathbf{b}$, and $\mathbf{c}$, we can immediately obtain the corresponding vertices $\mathbf{v}_1$, $\mathbf{v}_2$, and $\mathbf{v}_3$ (in 3D space) of the associated wave facet:

$$\mathbf{v}_1 = \mathbf{a} + z_a\mathbf{z} \tag{B.15a}$$

$$\mathbf{v}_2 = \mathbf{b} + z_b\mathbf{z} \tag{B.15b}$$

$$\mathbf{v}_3 = \mathbf{c} + z_c\mathbf{z}\,, \tag{B.15c}$$

where z_a, z_b, and z_c are the elevations of the sea surface at triad nodes $\mathbf{a}$, $\mathbf{b}$, and $\mathbf{c}$. These z values are determined as described in §3.2.1. As always, $\mathbf{z} = (0, 0, 1)$ is the unit vector normal to the mean sea surface.

Let $F(\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3)$ denote the wave facet defined by Eq. (B.15). A unit normal to this facet is given by

$$\mathbf{n}_1 = \frac{(\mathbf{v}_3 - \mathbf{v}_1) \times (\mathbf{v}_2 - \mathbf{v}_1)}{\|(\mathbf{v}_3 - \mathbf{v}_1) \times (\mathbf{v}_2 - \mathbf{v}_1)\|}\,,$$

where $\times$ denotes the vector cross product and $\|\mathbf{v}\|$ is the magnitude of $\mathbf{v}$. The unit *outward* normal to the facet is such that $\mathbf{n} \cdot \mathbf{z} > 0$, and is obtained by

$$\mathbf{n} = \text{sign}[\mathbf{n}_1 \cdot \mathbf{z}]\mathbf{n}_1\,. \tag{B.16}$$

Consider a plane in space with unit outward normal $\mathbf{n}$ and let $\mathbf{v}$ be any point in the plane. A necessary and sufficient for another point $\mathbf{q}$ to be in the plane is

$$(\mathbf{q} - \mathbf{v}) \cdot \mathbf{n} = 0\,.$$

The condition that the same point $\mathbf{q}$ lies on the ray $\mathbf{p} + s\boldsymbol{\xi}$ is

$$\mathbf{q} = \mathbf{p} + s(\mathbf{q})\boldsymbol{\xi}\,,$$

where $s(\mathbf{q})$ is the distance along the ray from $\mathbf{p}$ to $\mathbf{q}$ in the plane. We can solve for $s(\mathbf{q})$ from

$$(\mathbf{q} - \mathbf{v}_1) \cdot \mathbf{n} = (\mathbf{p} + s(\mathbf{q})\boldsymbol{\xi}) \cdot \mathbf{n}$$

whence

$$s(\mathbf{q}) = \frac{(\mathbf{v}_1 - \mathbf{p}) \cdot \mathbf{n}}{\boldsymbol{\xi} \cdot \mathbf{n}}\,,$$

where we have arbitrarily chosen vertex $\mathbf{v}_1$ of the facet $F(\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3)$ for the point $\mathbf{v}$.

We are therefore guaranteed, by construction of $s(\mathbf{q})$, that a point $\mathbf{q}$ on a ray $\boldsymbol{\xi}$ lies somewhere in the plane determined by the three vertices of wave facet $F(\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3)$. Point $\mathbf{q}$ in particular lies in the subset of the plane bounded by the facet itself if and only if

$$s(j_1) < s(\mathbf{q})\boldsymbol{\xi} \cdot \mathbf{n} < s(j_2)\,, \tag{B.17}$$

where j_1 and j_2 label the two triad intercept points associated with the wave facet, and where $s(j)$ is defined as in Eq. (B.7).

We interpret and use Eq. (B.17) as follows. Given a ray emanating from point $\mathbf{p}$ in direction $\boldsymbol{\xi}$, the track $\mathbf{p}_h + s\boldsymbol{\xi}_h$ across the hexagonal grid produces the triad intercept points $\boldsymbol{\pi}_0, \boldsymbol{\pi}_1, ..., \boldsymbol{\pi}_m$. Each pair of intercepts $(\boldsymbol{\pi}_j, \boldsymbol{\pi}_{j+1})$, $j = 0, 1, ...m-1$, determines the vertices $\mathbf{a}$, $\mathbf{b}$, and $\mathbf{c}$ of the associated triad, and thus the wave facet $F(\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3)$. The TIPs $\boldsymbol{\pi}_j$ and $\boldsymbol{\pi}_{j+1}$ are at distance $s(j)$ and $s(j+1)$, respectively, from $\mathbf{p}_h$ along direction $\boldsymbol{\xi}_h$. If the ray $\mathbf{p} + s\boldsymbol{\xi}$ intercepts the facet $F(\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3)$ at $\mathbf{q}$, a distance $s(\mathbf{q})$ from $\mathbf{p}$ along $\boldsymbol{\xi}$, then the horizontal projection in the plane of the hexagonal grid of the distance $s(\mathbf{q})$ from $\mathbf{p}$ to $\mathbf{q}$, $s(\mathbf{q})\boldsymbol{\xi} \cdot \mathbf{n}$, must lie between the limits indicated in Eq. (B.17). If (B.17) is not satisfied, then the ray $\mathbf{p} + s\boldsymbol{\xi}$ does not intercept the facet.

Whether or not a ray intercepts a sea surface facet within the hexagonal domain can be determined by systematically checking, in the above manner, each pair of triad intercept points, namely $(\boldsymbol{\pi}_0, \boldsymbol{\pi}_1)$, $(\boldsymbol{\pi}_1, \boldsymbol{\pi}_2)$, ..., $(\boldsymbol{\pi}_{m-1}, \boldsymbol{\pi}_m)$. Either the ray will intercept one of the facets at a point $\mathbf{q}$, or the sequence of triad intercept pairs $(\boldsymbol{\pi}_j, \boldsymbol{\pi}_{j+1})$ will be exhausted without ever satisfying Eq. (B.17), in which case the ray leaves the hexagonal domain without intersecting the sea surface.

If the ray $\boldsymbol{\xi}$ does intersect a facet, then the incident direction $\boldsymbol{\xi}$ and the facet normal $\mathbf{n}$ are used to compute the incident, reflected, and refracted directions for use in the Fresnel equations as described in §3.2.2. If ray $\boldsymbol{\xi}$ does not intersect a facet, then the ray information is tallied to the appropriate quad.

References

C. Bohren and E. Clothiaux. *Fundamentals of Atmospheric Radiation*. Wylie-VCH, 2006.

C. Bohren and D. Huffman. *Absorption and Scattering of Light by Small Particles*. Wyley, 1983.

M. Born and E. Wolf. *Principles of Optics, Fifth Edition*. Pergamon, 1975.

R. R. Bracewell. *The Fourier Transform and Its Applications, Second Edition, Revised*. McGraw-Hill, 1986.

S. Chandrasekhar. *Radiative Transfer*. Dover, 1960.

C. Cox and W. Munk. Measurement of the roughness of the sea surface from photographs of the sun's glitter. *J. Opt. Soc. Amer.*, 44:838–850, 1954.

T. Elfouhaily, B. Chapron, K. Katsaros, and D. Vandemark. A unified directional spectrum for long and short wind-driven waves. *J. Geophys. Res.*, 102:15781–15796, 1997.

R. D. M. Garcia. Fresnel boundary and interface conditions for polarized radiative transfer in a multilayer medium. *J. Quant. Spectros. Rad. Trans.*, 113:306–307, 2012.

E. Hecht. *Optics, Second Edition*. Addison-Wesley, 1989.

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

J. W. Hovenier. Multiple scattering of polarized light in planetary atmospheres. *Astron. Astrophys.*, 13:7–29, 1971.

Y.-X. Hu, D. Winker, P. Yang, B. Baum, L. Poole, and L. Vann. Identification of cloud phase from PICASSO-CENA lidar depolarization: a multiple scattering sensitivity study. *J. Quant. Spectros. Rad. Trans.*, 70:569–579, 2001.

G. W. Kattawar. Selected papers on multiple scattering in plane parallel atmospheres and oceans: Methods. In G. W. Kattawar, editor, *SPIE Milestone Series 42*, page 641. SPIE Optical Engineering Press, 1991.

G. W. Kattawar. Polarization of light in the ocean. In R. W. Spinrad, K. L. Carder, and M. J. Perry, editors, *Ocean Optics*, pages 200–225. Oxford Univ. Press, 1994.

G. W. Kattawar and C. N. Adams. Stokes vector calculations of the submarine light field in an atmosphere-ocean with scattering according to a Rayleigh phase matrix: Effect of interface refractive index on radiance and polarization. *Limnol. Oceanogr.*, 34:1453–1472, 1989.

M. I. Mischenko. Multiple scattering, radiative transfer, and weak localization in discrete random media: unified microphysical approach. *Rev. Geophys.*, 46:1–33, 2008.

M. I. Mischenko, L. D. Travis, and A. A. Lacis. *Scattering, Absorption, and Emission of Light by Small Particles*. Cambridge Univ. Press, 2002.

C. D. Mobley. *Light and Water: Radiative Transfer in Natural Waters*. Academic Press, 1994.

R. W. Preisendorfer. *Hydrologic Optics*. U.S. Department of Commerce, Pacific Marine Envoronmental Laboratory, 1976. In 6 volumes. Available online at *www.oceanopticsbook.info/view/introduction/level_2/text_books_relevant_to_ocean_optics*.

R. W. Preisendorfer and C. D. Mobley. Unpolarized irradiance reflectances and glitter patterns of random capillary waves on lakes and seas, by Monte Carlo simulation. Technical report, NOAA Pacific Marine Environmental Lab, 1985. Tech. Memo. ERL PMEL-63.

J. R. Schott. *Fundamentals of Polarimetric Remote Sensing*. SPIE Tutorial Text Vol. TT81, 1999.

Ø. Svensen, J. J. Stamnes, M. Kildemo, L. M. S. Aas, S. R. Erga, and Ø. Frette. Mueller matrix measurements of algae with different shape andd size distributions. *Appl. Optics*, 50:5149–5157, 2011.

Ø. Svensen, M. Kildemo, J. Maria, J. J. Stamnes, and Ø. Frette. Mueller matrix measurements and modeling pertaining to spectralon white reflectance standards. *Optics Exp.*, 20:15045–15053, 2012.

J. Tessendorf. Simulating ocean water. Technical report, SIGGRAPH Course Notes, 2004. URL `http://people.clemson.edu/~jtessen/reports.html`.

H. C. van de Hulst. *Multiple Light Scattering: Tables, Formulas, and Applications*. Academic Press, 1980.

K. J. Voss and E. S. Fry. Measurment of the Mueller matrix for ocean water. *Appl. Optics*, 23: 4427–4439, 1984.

C. R. Zeisse. Radiance of the ocean horizon. *J. Opt. Soc. Amer.*, 12:2022–2030, 1995.

P.-W. Zhai, G. W. Kattawar, and P. Yang. Impulse response solution of the three-dimensional vector radiative transfer equation in atmosphere-ocean systems. I. Monte Carlo method. *Appl. Optics*, 47:1037–1047, 2008.

P.-W. Zhai, Y. Hu, J. Chowdhary, C. R. Trepte, P. L. Luker, and D. B. Josset. A vector radiative transfer model for coupled atmosphere and ocean systems with a rough interface. *J. Quant. Spectros. Rad. Trans.*, 111:1025–1040, 2010.

P.-W. Zhai, G. W. Kattawar, and Y. Hu. Comment on the transmission matrix for a dielectric interface. *J. Quant. Spectros. Rad. Trans.*, 113:1981–1984, 2012.