On polarized light propagation in biological tissues.

Jaro Rička
Institute of Applied Physics, Sidlerstr. 5, 3012 Bern, Switzerland

May 14, 2009
# Contents

1 Introduction 3

2 Light as electromagnetic wave 4
   2.1 Maxwell Equations 4
   2.2 Plane waves in homogeneous media 5
      2.2.1 Anisotropic media 8
      2.2.2 Chiral media 10
   2.3 Beyond plane waves 10
      2.3.1 Beam-waves and their properties 10
      2.3.2 The two regimes: quasi-plane waves and quasi-spherical waves 12
      2.3.3 Light sources, dipole fields 14

3 Manipulating polarization 16
   3.1 Linear and circular polarization 16
   3.2 Basic hardware elements 17
      3.2.1 Retarder plate 17
      3.2.2 Linear polarizer 19
   3.3 Jones formalism 19
   3.4 Optical elements in Jones language 20
      3.4.1 Projecting filters 21
      3.4.2 Transformers 21
      3.4.3 Producing circular polarization 22
      3.4.4 Blocking and filtering circular polarization 23
   3.5 Deflectors 23
      3.5.1 Snellius and Fresnel laws in Jones language 24

4 Measuring polarization 27
   4.1 Basic concepts 27
   4.2 Time averaging, Stokes parameters 28
   4.3 General experiment design, Coherence matrix 29
   4.4 Stokes vectors 30

5 Beam filters and beam states 32
   5.1 Mathematical representation of beam filters 33
   5.2 Virtual beams as base vectors 35

6 Polarization analysis of an unknown optical system 36
   6.1 Designing the experiment, from Jones to Perrin-Mueller matrix 36
   6.2 Interpreting the Perrin-Mueller matrix, alternative codings 38
   6.3 Mueller versus Jones 39
Chapter 1

Introduction

The aim of the present “lecture notes” is to combine the photon-particle picture of light propagation and interaction in biological tissues with the electrodynamic model of light. In the classical electrodynamics light is regarded as a rapidly oscillating vector field. The term “light polarization” merely expresses the vector nature of the electromagnetic field. Two field vectors, $\mathbf{E}$ and $\mathbf{B}$, represent the forces experienced by charged matter. Recall the well known expression $\mathbf{F} = q(\mathbf{E} + \mathbf{v} \times \mathbf{B})$ for the Lorentz force acting on a free point-like charged particle. In biological tissue the magnetic forces turn out to be negligible and thus we shall be mainly concerned with the electric field $\mathbf{E}$. Because of the rapidity of light oscillations and because the charges are bound in atoms and molecules, there is no net motion of the charges. The only effect of the imposed electric field is the deformation of the charge system.

Concerning the interaction of laser radiation with the tissue, we can distinguish two extreme regimes: At low laser intensities, corresponding to low field strength $E$, we are in the linear regime where the charge displacements are proportional to $E$. Of course some of the energy of the laser radiation may be absorbed, causing heating of the tissue, photochemical reactions or re-emission of luminescence. In the low intensity regime, however, the absorption of the laser radiation can be expected to have only negligible effect on the optical properties of the tissue. This is the regime of optical diagnostics and certain minimally invasive therapeutic methods, such as photodynamic therapy. In this context we are mostly interested in the light propagation through the tissue, here specifically in the effect of the polarization on the light propagation. The other extreme is the highly non-linear regime of laser ablation, but in this regime polarization plays only a minor role. Thus, the present lecture notes are aimed at the diagnostic applications of laser radiation. A topic of growing interest in optical tissue diagnostics, especially in dermatology, is polarization imaging [1, 2, 3]. The first step in developing novel diagnostic methods is usually the comparison: one compares the Mueller matrix images of a healthy and diseased tissue, searching for a signature of the disease. In a second step one would like to understand how such images are formed in order to link the signature with changes of the micro-structure of the tissue that are associated with the disease. This is a formidable task, considering the complexity of the biological tissue as an optical medium.

The lecture notes are organized as follows: In Chapters 2-6 we compile concepts and tools that we deem to be necessary or convenient for understanding the propagation of polarized light and for polarization analysis. In Chapters 7-9 these tools will be put into operation for the development of the concepts for realistic optical tissue modeling and finally in Chapters 10-11 we discuss the concepts of radiative transfer and polarized Monte Carlo simulations.
Chapter 2

Light as electromagnetic wave

2.1 Maxwell Equations

The topic of electromagnetism is the subject of numerous standard books (e.g. [4, 5]). We only need to review the matters in a language appropriate for the description of light propagation in biological tissue. (Thereby we shall employ SI-units for electromagnetic quantities, which include the vacuum permittivity $\varepsilon_0\left[\text{CV}^{-1}\text{m}^{-1}\right]$ and vacuum permeability $\mu_0\left[\text{VsA}^{-1}\text{m}^{-1}\right]$. Recall that $\varepsilon_0\mu_0 = 1/c^2$, where $c$ is the speed of light in vacuum.) The core of the electromagnetic theory is the celebrated set of Maxwell Equations (ME) relating the field vectors $E(r,t)\left[\text{V/m}\right]$ and $B(r,t)\left[\text{Vs/m}\right]$ with their sources, namely with the charge distribution $\rho(r,t)\left[\text{C/m}^3\right]$ and current distribution $j(r,t)\left[\text{A/m}^2\right]$. Employing the SI units, the Maxwell Equations read:

\begin{align*}
\nabla \times E + \frac{\partial B}{\partial t} &= 0 \quad (2.1) \\
\frac{1}{\mu_0} \nabla \times B - \varepsilon_0 \frac{\partial E}{\partial t} &= j \quad (2.2) \\
\varepsilon_0 \nabla \cdot E &= \rho \quad (2.3) \\
\nabla \cdot B &= 0 \quad (2.4)
\end{align*}

Macroscopic currents and free charges play hardly any role in the interaction of the rapidly oscillating light field with condensed matter. Therefore we only consider the microscopic contributions to $\rho$ and $j$, which are due to charges bound in atoms or molecules. These contributions originate chiefly from electric dipoles $d_m$ and magnetic dipoles $m_m$ that are induced in the molecules by the applied fields. Correspondingly, the charge and current distributions $\rho$ and $j$ reflect the distributions $P$ and $M$ of the molecular electric and magnetic dipoles (traditionally called “dielectric polarization” and “magnetization”). In a simple microscopic model one would express these distributions in terms of Dirac’s δ-functions representing the position of the molecules[6], as for example $P(r) = \sum_m d_m \delta(r-r_m)$. We shall need such microscopic approach for the development of light scattering in tissues, but for now we neglect the molecular picture and write the bound charges and currents in the standard form ([4, 5]) as

\begin{align*}
\rho &= -\nabla \cdot P \quad j = \frac{\partial P}{\partial t} + \nabla \times M \quad (2.5)
\end{align*}

Note that the equation system 2.1-2.5 is not yet closed. Since $P$ and $M$ have been generated by the fields $E$ and $B$, we will have to supply the constitutive relations $P = P(E,B)$ and $M = M(E,B)$. To simplify the matters, we restrict ourselves to
monochromatic radiation. This is legitimate because laser light is in a good approximation monochromatic, and because any time dependence can be synthesized as a superposition of appropriately phased monochromatic waves. Furthermore we assume that the response of the matter is linear. Thus all field vectors and molecular dipole moments oscillate with the same angular frequency $\omega$. For example

$$\mathbf{P}(\mathbf{r}, t) = \mathbf{P}(\mathbf{r}) e^{-i\omega t} \quad \text{and} \quad \mathbf{E}(\mathbf{r}, t) = \mathbf{E}(\mathbf{r}) e^{-i\omega t}$$

(2.6)

Note that the vectors are complex quantities, for example $\mathbf{E} = \Re \mathbf{E} + i \Im \mathbf{E}$. The classical fields are real parts. The complex notation simplifies the mathematics, because the differential operator $\partial / \partial t$ is replaced with multiplication by $-i\omega$. But this is not the only motivation for the complex notation. We intend to combine the photon picture with the electrodynamic model of light and this requires to accept at least some of the concepts of the quantum theory. In the quantum world we deal quite naturally with complex wave functions, while the observables are real.

The observable in which we are most interested in the context of light propagation is the photon flux density $N$ [m$^{-2}$] or the irradiance $I$ [Wm$^{-2}$] (symbol $I$ is used in the present chapter to avoid confusion with the electric field strength $E$). In the classical electrodynamics the energy flow is represented by the Poynting vector $\mathbf{S}(t) = \left[ \Re \mathbf{E}(t) \times \Re \mathbf{B}(t) \right] / \mu_0$. However, since light fields exhibit extremely rapid oscillations, the only meaningful observable is the average $\mathbf{S} = \overline{\mathbf{S}}(t)$ over at least a couple of oscillation cycles. In the complex notation this quantity is written as\footnote{In most textbooks one finds Poynting vector defined as $\mathbf{S}' = \frac{1}{2} \Re \left[ \mathbf{E}^* \times \mathbf{H} \right]$ where $\mathbf{H}$ is the auxiliary fields defined in Eqs 2.9. The discussion whether $\mathbf{S}'$ or $\mathbf{S}$ is the correct expression for the electromagnetic Poynting vector in condensed matter is continuing [7, 8] and likely to be intensified in the context of the so-called “lethanded metamaterials” [9]. Luckily, magnetization in biological tissues at optical frequencies is negligible, so that $\mathbf{S}' = \mathbf{S}$.}

$$\mathbf{S} = S \hat{s} = \frac{1}{4 \mu_0} \left[ \mathbf{E} \times \mathbf{B}^* + \mathbf{E}^* \times \mathbf{B} \right] = \frac{1}{2 \mu_0} \Re \left[ \mathbf{E}^* \times \mathbf{B} \right]$$

(2.7)

The power $dP$ flowing through an oriented infinitesimal area $d\mathbf{a} = da \hat{n}$ is given by $dP = \mathbf{S} d\mathbf{a} = S(\hat{s} \cdot \hat{n}) da$, where $\hat{s}$ is the unit vector pointing in the direction of $\mathbf{S}$ and $\hat{n}$ is the unit vector specifying the orientation of the surface. To highlight the significance of the Poynting vector, we recall the Poynting’s theorem [4, 5] for harmonic fields:

$$\nabla \cdot \mathbf{S} = -\frac{1}{2} \Re (j \cdot \mathbf{E}^*) .$$

(2.8)

Equation 2.8, which follows straightforwardly from Maxwell Equations, expresses the conservation of electromagnetic energy: divergence of electromagnetic energy flux (left hand side) is due to rate of exchange of electromagnetic energy with the matter, e.g. heating by absorption (right hand side). One should keep in mind that $\mathbf{E}$ and $\mathbf{B}$ are complicated microscopic fields that consist of the superposition of the field of the driving laser and of the fields whose sources are the oscillating charges in the matter. One can think of the photons as being engaged in nearly endless microscopic interaction loops, which as we know, slow down their propagation.

2.2 Plane waves in homogeneous media

The exact solution of the Maxwell Equations in a randomly heterogeneous medium such as a biological tissue is an impossible task. We are forced to make approximations.
Useful zero order approximation is to replace the tissue by a homogeneous effective medium, whose molecular constituents are smoothed by a suitable averaging procedure [6]. Then, one regards the dipole distributions \( P \) and \( M \) as smooth vector fields like \( E \) and \( B \), which may be included in two auxiliary fields \( D \) and \( H \):

\[
D = \varepsilon_0 E + P \quad \quad H = \frac{1}{\mu_0} B - M. \tag{2.9}
\]

Here \( D \) is the so-called “dielectric displacement”, which incorporates the dielectric response of the matter. The magnetic response of the matter is incorporated in \( H \). However, biological tissues do not exhibit magnetic response at optical frequencies and so we will usually set \( M = 0 \). (Magnetic dipoles will return through a back door of chiral, optically active media.)

The smoothed version of Maxwell Equation is particularly convenient for an approximate yet self-consistent treatment of electromagnetic waves in condensed matter. Recalling from the elementary scalar optics the role of the harmonic plane waves, we make the following *ansatz*: space and time dependence of the field vectors are expressed by common oscillating phase factor \( \exp(ik \cdot r - i\omega t) \), where \( \omega \) is the oscillation frequency and \( k \) is the wave vector. In general, \( k \) is a complex vector, whose imaginary part indicates an attenuated or an evanescent wave. In a non absorbing medium \( k \) is a real quantity, the unit vector \( \hat{k} \) is perpendicular to the wave front and the magnitude \( k = |k| \) defines the phase velocity of light in the given direction: \( v = \omega/k \). The same phase factor also applies to the material response, as long as we remain in the linear regime.

Thus we write for example \( E(r,t) = E \exp(ik \cdot r - i\omega t) = E\hat{e} \exp(ik \cdot r - i\omega t) \), where \( E \) is a constant complex vector, \( E \) is the scalar amplitude and \( \hat{e} \) is a complex unit vector such that \( \hat{e}^* \cdot \hat{e} = 1 \). This vector \( \hat{e} \) specifies the polarization state of the wave, as shall be discussed in detail in Sec. 3. For now we only note that \( \hat{e} \) is a real unit vector in the case of a linearly polarized wave. Note that this \( E(r,t) \) is not yet the solution of ME, it is only a convenient *ansatz*: in the \( k,\omega \)-representation the differential operators are transformed into simple vector operations \( \partial/\partial t \rightarrow -i\omega \) and \( \nabla \rightarrow ik \), and the common phasor \( \exp(ik \cdot r - i\omega t) \) can be dropped from all equations. This reduces the set of Maxwell Equations into an exercise in vector algebra:

\[
\begin{align*}
k \times E &= \omega B \\
\quad & \quad \tag{2.10} \\
\quad & \quad \\
k \times H &= -\omega D \\
\quad & \quad \tag{2.11} \\
\quad & \quad \\
k \cdot D &= 0 \\
\quad & \quad \tag{2.12} \\
\quad & \quad \\
k \cdot B &= 0 \\
\quad & \quad \tag{2.13}
\end{align*}
\]

We also recall Eqs. 2.9, still assuming that the medium is non magnetic, i.e. \( M = 0 \). Upon examining this set of equations, one discovers certain geometrical relationships, which are summarized in Fig. 2.1, assuming thereby that \( k \) is real: from 2.12 and 2.13 it follows that \( k \) is normal to a plane spanned by \( B \) and \( D \). In addition, according to 2.11, \( D \) is normal to a plane spanned by \( k \) and \( H = B/\mu_0 \) (which points into the paper.) Thus, \( B, D \) and \( k \) form an orthogonal tripod. From 2.10 we deduce that \( B \) is normal to the plane \( k, E \). Thus \( B \) and \( E \) are orthogonal and \( E \) must lay in the plane \( k, D \). However the angle \( \gamma \) between \( E \) and \( D = \varepsilon_0 E + P \) is not yet specified, since it depends on the angle between \( P \) and \( E \). Energy propagates in the direction of the Poynting vector (recall Eq. 2.7), i.e. perpendicularly to \( E \) and \( B \). Note that \( S \) and \( k \) are not necessarily parallel but they span the same angle \( \gamma \) as \( E \) and \( D \). In any case, the vectors \( E, D, P, k \) and \( S \) lay in a plane to which \( B \) is normal. The same considerations apply to the real and imaginary parts of the complex field vectors.
Figure 2.1: Geometrical relationships in a harmonic plane wave. Magnetic vector $\mathbf{B}$ points into the paper. Note that the Poynting vector $\mathbf{S}$ and the wave vector $\mathbf{k}$ are not necessarily parallel. The light gray lines represent the wave fronts, which are per definition normal to $\mathbf{k}$. They are parallel with $\mathbf{D}$, but not necessarily with $\mathbf{E}$. In an isotropic medium one would have $\mathbf{D} \parallel \mathbf{E}$ and therefore $\mathbf{S} \parallel \mathbf{k}$. (The outlined situation illustrates the propagation of an extraordinary wave in an anisotropic medium, $\mathbf{\hat{a}}$ is the anisotropy axis, grayly printed quantities are explained in Sec. 2.2.1.)

After the exercise in geometry, all we have to do is to solve the Maxwell Equations for the unknown magnitude $k$ of the wave vector and for the angle $\gamma$. First we eliminate $\mathbf{H} = \mathbf{B}/\mu_0$ from Eq. 2.11, exploiting thereby Eq. 2.10. Using then a well known vector identity $\mathbf{k} \times (\mathbf{k} \times \mathbf{E}) = \mathbf{k} \cdot (\mathbf{k} \cdot \mathbf{E}) - (\mathbf{k} \cdot \mathbf{k}) \mathbf{E}$ we arrive at a vectorial wave equation in the $\mathbf{k}, \omega$-representation:

$$(\mathbf{k} \cdot \mathbf{k}) \mathbf{E} - \mathbf{k} (\mathbf{k} \cdot \mathbf{E}) = k_o^2 (\mathbf{E} + \mathbf{P}/\epsilon_o)$$

(2.14)

Here $k_o = \omega/c$ is the wave number in vacuum. In order to solve Eq. 2.14, we must provide a constitutive relation $\mathbf{P} = \mathbf{P}(\mathbf{E})$ appropriate for the medium under consideration. The general solution is a little bit tricky and therefore we first exercise using the familiar case of an isotropic dielectrics:

$$\mathbf{P} = \epsilon_o \chi \mathbf{E} \quad \mathbf{D} = \epsilon_o (1 + \chi) \mathbf{E} = \epsilon \epsilon_o \mathbf{E}$$

(2.15)

where $\chi$ is the susceptibility and $\epsilon = 1 + \chi$ is the relative permittivity. Combining the isotropic constitutive relation with Eq. 2.12 one finds that $\mathbf{k} \cdot \mathbf{E} = 0$. (In other words, $\mathbf{k}$ and $\mathbf{E}$ are perpendicular, which also implies that $\mathbf{\hat{s}} \parallel \mathbf{k}$. In isotropic media energy is transported along $\mathbf{k}$.) Thus, Eq. 2.16 reduces to the identity

$$n^2 \mathbf{E} = \epsilon \mathbf{E}, \quad \text{where} \quad n^2 = \frac{k \cdot k}{k_o^2}.$$  

(2.16)

which is fulfilled if $n^2 = (k \cdot k)/k_o^2 = \epsilon$. The right hand side of Eq. 2.16 is the general definition of refractive index $n$, applicable for real or complex $k$. (Note that if $k$ is a complex vector then $k \cdot k$ is not the same as $|k|^2 = k \cdot k^\ast$.) Assuming for simplicity that $\epsilon$ and $k$ are real (no losses, no evanescence), we obtain the familiar result $k \cdot k = k^2 = n^2 k_o^2$, where $n = \sqrt{\epsilon}$ is the refractive index. We leave it to the reader to verify that Eq. 2.16 is nothing but a disguise of the familiar wave equation $c_m^2 \nabla^2 \mathbf{E} = \partial^2 \mathbf{E}/\partial t^2$, where $c_m^2 = c^2/\epsilon$ (recall that $\partial/\partial t \rightarrow -i\omega$ and $\nabla \rightarrow ik$).
2.2.1 Anisotropic media

Having rehearsed the familiar concepts, we tackle a more general case. Biological tissues are known to exhibit a certain degree of structural anisotropy. They contain elongated fiber-like macromolecules, and those fibers may be organized in oriented structures resulting in structural birefringence. A second not less important motivation is that birefringent optical elements are needed for measuring and manipulating the polarization of light. The anisotropy in structure results in the anisotropy of dielectric response of the material to the applied field: along certain directions the charges respond more or less susceptibly to the electric force \( E \). We restrict ourselves to the most common case of uniaxial anisotropy: there is a single special direction, a single anisotropy axis \( \hat{a} \). The uniaxial constitutive relation reads

\[
P/\varepsilon_0 = \chi E + \Delta \chi \hat{a} (\hat{a} \cdot E)
\]  

(2.17)

Here \( \chi \) is the overall isotropic contribution to the dielectric susceptibility and \( \Delta \chi \) is the anisotropic susceptibility excess (positive or negative) in the direction \( \hat{a} \). The isotropic case is recovered by setting \( \Delta \chi = 0 \). Note that the vector \( E_{\hat{a}} = \hat{a} (\hat{a} \cdot E) \) represents the projection of the \( E \) vector on the anisotropy axis. Often Eq. 2.17 can be found in the short hand tensor notation \( \mathbf{P} = \varepsilon_0 \chi \mathbf{E} \), where the components of \( \chi \) are \( \chi_{kl} = \chi \delta_{kl} + \Delta \chi a_k a_l \). (Adding one more anisotropy axis we would naturally get the biaxial case, but adding still more would not bring anything new, because a 3D tensor has only 3 diagonal components. Note that in the uniaxial case the tensor \( a_k a_l \) is a simple projection operator.) We prefer here the vector notation, since it makes it easier to understand the geometrical relationships. Further we assume for simplicity a lossless medium (\( k \) real) and linear polarization (\( \hat{e} \) real). Inserting the Eq. 2.17 into 2.14 and dividing by the common factor \( |k|/E \), the wave equations for light propagation in anisotropic media become

\[
\hat{e} - \hat{k} \left( \hat{k} \cdot \hat{e} \right) = \frac{\varepsilon}{n^2} \hat{e} + \frac{\Delta \chi}{n^2} \hat{a} (\hat{a} \cdot \hat{e})
\]  

(2.18)

\[
\left( \hat{k} \cdot \hat{e} \right) = -\frac{\Delta \chi}{\varepsilon} \left( \hat{k} \cdot \hat{a} \right) (\hat{a} \cdot \hat{e}).
\]  

(2.19)

The second part follows from Eq. 2.12. The task can be formulated as follows: Suppose that in a medium with anisotropy axis \( \hat{a} \) we would like to excite a plane wave whose wave front propagates in the direction \( \hat{k} \), as indicated in Fig. 2.1. We know the inclination cosine \( \hat{k} \cdot \hat{a} = \cos(\beta) \) and the medium parameters \( \varepsilon, \Delta \chi \). What are the permitted values of \( |k|^2 = n^2 k_0^2 \) and what polarization directions \( \hat{e} \) are allowed? Note that the questions can be answered quite easily if \( \hat{k} \cdot \hat{a} = 1 \) or if \( \hat{k} \cdot \hat{a} = 0 \). The letter special case, with optical axis \( \hat{a} \) oriented in right angle with \( \hat{k} \), is realized in polarizing optical elements such as retarders or Glan Thompson prism. However, in birefringent biological tissues light propagates in a random direction and therefore we need the general results. Upon examining the geometry of the problem, one finds that the sought information on \( \hat{e} \) is contained in two direction cosines \( \hat{a} \cdot \hat{e} \) and \( \hat{k} \cdot \hat{e} \). Thus, we convert the vector equation 2.18 into two scalar equations by multiplying both sides with \( \hat{a} \) and \( \hat{e} \), which gives a

\footnote{Anisotropy of dielectric response should not be confused with anisotropic scattering from large particles. There is no macroscopic scattering in homogeneous media.}
Because of ordinary plane wave in an isotropic medium, therefore it is called the “ordinary” wave.

\[ (\hat{a} \cdot \hat{e}) - (\hat{a} \cdot \hat{k}) (\hat{k} \cdot \hat{e}) = \frac{\epsilon}{n^2} (\hat{a} \cdot \hat{e}) + \frac{\Delta \chi}{n^2} (\hat{a} \cdot \hat{e}) \]  
\[ 1 - (\hat{k} \cdot \hat{e})^2 = \frac{\epsilon}{n^2} + \frac{\Delta \chi}{n^2} (\hat{a} \cdot \hat{e})^2 \]
\[ (\hat{k} \cdot \hat{e}) = -\frac{\Delta \chi}{\epsilon} (\hat{k} \cdot \hat{a}) (\hat{a} \cdot \hat{e}) \]

An algebra program (Maple or Mathematica) suggests two solutions, which is not too surprising, since we are dealing with the phenomenon of birefringence. The so-called “ordinary” solution could have been already guessed when looking at Eq. 2.18:

\[ \hat{a} \cdot \hat{e}_\perp = 0 \quad \hat{k} \cdot \hat{e}_\perp = 0 \quad n^2 = \epsilon \]  

Apparently, the polarization vector \( \hat{e}_\perp \) is perpendicular to both, \( \hat{a} \) and \( \hat{k} \), and therefore perpendicular to the plane spanned by \( \hat{a} \) and \( \hat{k} \). The electric field oscillates perpendicularly to the propagation direction \( \hat{k} \), which then coincides with the direction \( \hat{s} = S/|S| = \hat{e} \times \hat{b} \) of energy propagation (recall Eq. 2.7). Such a wave behaves just as an ordinary plane wave in an isotropic medium; therefore it is called the “ordinary” wave. Because of \( \hat{e}_\perp \parallel \hat{a} \), the ordinary wave does not feel the anisotropy, i.e., it experiences the refractive index \( n = \sqrt{\epsilon} \).

The second solution is the “extraordinary” wave, illustrated in Fig. 2.1:

\[ (\hat{a} \cdot \hat{e}_\parallel)^2 = \frac{\sin(\beta)^2}{\sin(\beta)^2 + \cos(\beta)^2 (\epsilon + \Delta \chi)^2 / \epsilon^2} \]
\[ (\hat{k} \cdot \hat{e}_\parallel) = -\frac{\Delta \chi}{\epsilon} \cos(\beta) (\hat{a} \cdot \hat{e}_\parallel) \]

One verifies, that the projection of \( \hat{e}_\parallel \) on the \( ak \)-plane, namely the vector \( e_{ak} = \hat{k} (\hat{k} \cdot \hat{e}_\parallel) + \hat{a} (\hat{a} \cdot \hat{e}_\parallel) \), is a unit vector. Thus, \( \hat{e}_\parallel = e_{ak} \) lays in the \( ak \)-plane. (Note that \( \hat{e}_\parallel \) and \( \hat{e}_\perp \) are orthogonal.) The result for the extraordinary refractive index is:

\[ \frac{1}{n^2} = \frac{\cos(\beta)^2}{\epsilon} + \frac{\sin(\beta)^2}{\epsilon + \Delta \chi} \]  

Note that Eq. 2.23 represents an ellipse. Refractive index experienced by the extraordinary wave depends on the direction \( k \) of the wave propagation with respect to the anisotropy axis. In the special case \( \hat{k} \cdot \hat{a} = \cos(\beta) = 1 \) one recovers the ordinary wave with \( \hat{a} \cdot \hat{e} = 0 \) and \( n = \sqrt{\epsilon} \). This is because the \( \hat{a} \)-component of \( \hat{e}_\parallel \) vanishes and thus the wave does not feel the anisotropy. On the other hand, when \( \hat{k} \perp \hat{a} \), then the wave experiences refractive index \( \sqrt{\epsilon + \Delta \chi} \). If \( \Delta \chi > 0 \) (positive birefringence) then wave propagation is slowed down with respect to the ordinary wave. (Thus, the “slow axis” of a retarder plate made from a positively birefringent material coincides with \( \hat{a} \), see Sec. 3.2.1). The truly extraordinary feature of the extraordinary wave is the fact that the direction \( \hat{s} = S/|S| \) of energy propagation does not coincide with the direction \( \hat{k} \) of wave propagation. Evaluating the extraordinary Poynting vector (combine Eq. 2.7 with Eq. 2.10), one obtains:

\[ S_{ext} = |k| \left[ \hat{k} - (\hat{k} \cdot \hat{e}_\parallel) \hat{e}_\parallel \right] \frac{1}{2} \epsilon_o |E|^2 c, \]

from where it follows:

\[ \hat{s} = \left[ \hat{k} - (\hat{k} \cdot \hat{e}_\parallel) \hat{e}_\parallel \right] / \sqrt{1 - (\hat{k} \cdot \hat{e}_\parallel)^2} \quad \hat{k} \cdot \hat{s} = \cos(\gamma) = \sqrt{1 - (\hat{k} \cdot \hat{e}_\parallel)^2} \]
Both, the ordinary and the extraordinary wave, are simultaneously excited when the anisotropic material is illuminated with a beam whose polarization is neither fully parallel nor fully perpendicular to the plane spanned by \( \mathbf{a} \) and \( \mathbf{k} \). This splits the beam into two beams with orthogonal polarizations \( \hat{e}_\parallel \) and \( \hat{e}_\perp \): one observes birefringence. Because of the wave splitting, we were allowed to treat the ordinary wave \( \mathbf{E}_\perp \) and the extraordinary wave \( \mathbf{E}_\parallel \) separately.

### 2.2.2 Chiral media

Another special type of the material response often found in biological tissues is chiral response, which manifests itself as optical activity, i.e. rotation of the polarization plane of linearly polarized light. A well known example is the optical activity of glucose solution, which suggests to use the effect for non-invasive measurement of sugar levels in diabetes patients (see, however, Ref. [10]). The microscopic origins of optical activity are chiral, i.e. screw-like molecules (such as glucose), whose structure forces the charges to move in a screw-like fashion. Thus, the electric force generates in the medium microscopic coil magnets and we end up with electrically induced magnetic dipole density \( \mathbf{M}_e \). Since \( \mathbf{E} \) accelerates charges, we expect \( \partial \mathbf{M}_e / \partial t \sim \mathbf{E} \), which leads to the chiral wave equation [5]:

\[
(k^2 - k_o^2 \epsilon) \mathbf{E} = -i k \xi \hat{\mathbf{k}} \times \mathbf{E} \quad (2.26)
\]

where \( \epsilon = 1 + \chi \) and \( \xi \) are dimensionless material constants. For the solution we use the technique of educated guess. Realizing that the cross-product \( \hat{\mathbf{k}} \times \mathbf{E} \) actually means rotation, we guess that circular polarization states could solve the problem. There are two orthogonal circular states, represented by the complex unit vectors \( \hat{e}_L = 1/\sqrt{2} (\hat{x} - i \hat{y}) \) and \( \hat{e}_R = 1/\sqrt{2} (\hat{x} + i \hat{y}) \). The subscripts \( L \) and \( R \) refer to left- and right-circularly polarized light waves. Inserting these states into Eq. 2.26 we obtain

\[
(k^2 - k_o^2 \epsilon) \hat{e}_{L/R} = \pm k \xi \hat{e}_{L/R}, \quad (2.27)
\]

where + is for \( \hat{e}_L \) and - for \( \hat{e}_R \). This gives two simple quadratic equations for \( k \) to be solved. Choosing the solutions such that \( k > 0 \) (one propagation direction of the wave), one obtains

\[
k_{L/R} = \sqrt{k_o^2 \epsilon + \xi^2 / 2} \pm \xi / 2, \quad (2.28)
\]

where again + is for \( \hat{e}_L \) and - for \( \hat{e}_R \). Thus, we can write the refractive index as \( n_{L/R} = n \pm \delta n \), where \( n = \sqrt{\epsilon + \xi^2 / 2k_o^2} \) and \( \delta n = \xi / 2k_o \). In other words, the two circular polarizations propagate with different phase velocities, undergoing a mutual phase shift during their propagation. One consequence of the refractive index difference is circular birefringence: a light beam passing through a prism made from chiral material would split into two beams with mutually orthogonal circular polarizations. Another consequence of the phase shift is the optical activity: after passing a linearly polarized beam through a chiral slab, one finds the polarization plane rotated in a certain angle with respect to the input.

### 2.3 Beyond plane waves

#### 2.3.1 Beam-waves and their properties

In the previous sections we pretended to be working with photon states that are plane waves with perfectly defined wave vector \( \mathbf{k} \), but we could not avoid talking or thinking in terms of “beams” or “rays”. This is because there is a problem with the plane
waves: they may be legal solutions of Maxwell equations in homogeneous media, but they are only idealized mathematical constructs. They can’t exist in practice, because if the amplitude *E* of a plane wave is finite, then the plane wave transports infinite power. Plane wave photon states represent an extreme of the uncertainty relation: they are perfectly localized in the \( \textbf{k} \)-space, but they are completely delocalized in the real space. The probability to detect a photon is the same at any position \( \textbf{r} \). Consequently, plane waves are useless for imaging. Because polarization imaging is an increasingly important field of tissue optics, we must find better ways to represent photons. Before the wave nature of light was accepted, classical optics provided a sort of solution: geometrical rays. They are still a useful practical concept, but they are not a good representation of physical photon states: a ray is perfectly localized in the transverse direction and it also has a well defined direction of propagation. In other words, a ray is perfectly localized both in \( \textbf{k} \) and \( \textbf{r} \), violating thus the uncertainty relation in the most drastic way. Moreover, it is quite difficult to equip a ray with polarization and to include interference effects. Fortunately, there is a simple way out of the dilemma: there exist photon states which are fairly localized both in \( \textbf{r} \) and \( \textbf{k} \) (as needed for imaging), which resemble plane waves in many respects, and which occur quite naturally in every photonic laboratory. Those photon states are beams. A monochromatic beam as a spatially coherent superposition of plane waves\(^3\):

\[
E(\textbf{r}) = \frac{E}{2\pi} \int_{-1}^{1} \int_{-1}^{1} \hat{e}(\kappa_x, \kappa_y) B(\kappa_x, \kappa_y) e^{i\kappa \cdot \textbf{r}} d\kappa_x d\kappa_y, \tag{2.29}
\]

\[
B(\textbf{r}) = \frac{1}{c} \frac{E}{2\pi} \int_{-1}^{1} \int_{-1}^{1} \hat{\textbf{k}}(\kappa_x, \kappa_y) \times \hat{e}(\kappa_x, \kappa_y) B(\kappa_x, \kappa_y) e^{-i\kappa \cdot \textbf{r}} d\kappa_x d\kappa_y.
\]

Here, \( \hat{e} \) is the polarization vector and \( \hat{\textbf{k}} = \textbf{k}/k \) is the propagation unit vector with components \( \kappa_x, \kappa_y \) and \( \kappa_z \) such that \( \kappa_x^2 + \kappa_y^2 + \kappa_z^2 = 1 \). For simplicity we assume that the beam propagates along the \( z \)-axis and that the beam focus is at \( \textbf{r}_f = \textbf{0} \). The beam profile is determined by the distribution \( B(\kappa_x, \kappa_y) \) of the complex amplitudes of the partial waves. \( B(\kappa_x, \kappa_y) \) is quadratically normalized:

\[
\int \int |B(\kappa_x, \kappa_y)|^2 d\kappa_x d\kappa_y = 1. \tag{2.30}
\]

An important special case is a Gaussian beam, a good approximation for the fundamental mode of a laser resonator and the mode emanating from a single mode fiber:

\[
B(\kappa_x, \kappa_y) = \frac{1}{\sqrt{\pi} \sigma_k} e^{-(\kappa_x^2 + \kappa_y^2)/2\sigma_k^2} \tag{2.31}
\]

The profile of a Gaussian beam is completely characterized by two numbers: the wave number \( k \) and the variance of the radial component of \( \hat{\textbf{k}} \), i.e., the average \( \sigma_k^2 = \langle \kappa_x^2 + \kappa_y^2 \rangle = \int \int (\kappa_x^2 + \kappa_y^2) |B(\kappa_x, \kappa_y)|^2 d\kappa_x d\kappa_y \). In cylindrical coordinates one has \( \sigma_k^2 = \langle \sin^2(\beta_k) \rangle^2 \), where \( \beta_k \) is the angle of the inclination of the partial wave \( \hat{\textbf{k}} \) with respect to the beam axis \( z \). Thus, we identify \( \sigma_k = \sqrt{\langle \sin^2(\beta_k) \rangle} \) as the effective numerical aperture of a beam (of any beam, not only Gaussian). Note that if \( B(\kappa_x, \kappa_y) \) is sufficiently narrow then \( \sigma_k \approx \beta = \sqrt{\langle \beta_k^2 \rangle} \), where \( \beta \) is the divergence angle of the beam.

Unlike plane waves, beams with finite amplitude \( E \) carry finite power \( P \)

\[
P = \oint_A \mathbf{S} \cdot d\mathbf{a}, \tag{2.32}
\]

\(^3\)In a medium replace \( c \) with \( c_m = c/n \) and \( \epsilon_o \) with \( \epsilon_m = \epsilon\epsilon_o; \ k \) is the wave number in the medium.
where \( A \) is an arbitrary surface enclosing the source of the beam (i.e. the laser) and \( \mathbf{a} \) is the outer normal to this surface. Since lasers usually radiate in one direction, one can take for \( A \) an arbitrary plane intersected by the beam. Upon inserting the Poynting vector that follows from Eq. 2.29 into Eq. 2.32 and using Eq. 2.30 one obtains:

\[
P = \frac{c \epsilon_0}{2} E^2 \frac{1}{k^2} \langle \kappa_z \rangle,
\]

(2.33)

where \( \langle \kappa_z \rangle = \langle \cos(\beta_k) \rangle = \int \int |B(\kappa_x, \kappa_y)|^2 \kappa_z \, d\kappa_x \, d\kappa_y \) is the mean inclination cosine of the partial waves.

### 2.3.2 The two regimes: quasi-plane waves and quasi-spherical waves

Regard in Fig. 2.2 a sketch of the wave structure of a typical moderately focused beam. One clearly observes two distinct regimes: In the asymptotic regime, far from the focus, the wave fronts are spherical, light propagates as a quasi-spherical wave. The asymptotic version of Eq. (2.29) is ([11] and references therein):

\[
E(\mathbf{r})|_{R \to \infty} = E \mathbf{e}(s_x, s_y) B(s_x, s_y) s_z \frac{e^{ikR}}{kR}
\]

(2.34)

\[
B(\mathbf{r})|_{R \to \infty} = \frac{1}{c} \hat{s} \times E(\mathbf{r})|_{R \to \infty}
\]

Here, \( R = (x^2 + y^2 + z^2)^{1/2} \) is the distance from the beam focus at \( \mathbf{r}_0 = 0 \) and \( \hat{s} \) is the unit vector pointing to the direction of observation with the components \( s_x = x/R \), \( s_y = y/R \) and \( s_z = z/R \). The second part of Eq. (2.34) expresses the transversality of the field in the radiation zone, far from focus. The asymptotic field of Eq. (2.34) exhibits a remarkable property: the amplitude of the spherical wave \( \exp(i kR)/kR \) at a point \( \mathbf{r} = \hat{s} R \) in the far-field is determined by a single partial wave with amplitude \( B(\kappa_x, \kappa_y) \) and polarization \( \mathbf{e}(\kappa_x, \kappa_y) \), such that \( \kappa_x = s_x = x/R \) and \( \kappa_y = s_y = y/R \). In fact, each geometrical ray in the far field represents a partial wave. Hence, probing the field at some point far from the focus in the direction of propagation, one picks up a single partial plane wave, weighted with the factor \( s_z = z/R \). On the other hand, one can go into the far field against the direction of propagation, before the focus, and do there all sorts of beam engineering. For example, by strongly focusing a beam that is radially polarized in the far field, one can even create axial polarization in the focus [12], which may be useful for laser ablation.

![Figure 2.2: Schematics of the wave structure of a moderately focused Gaussian beam. The contours indicate intensity levels. \( a_w \) is the 1/e waist radius, \( z_r = k a_w^2 = 1/k \sigma_k^2 \) is the size of the Rayleigh range. For paraxial beams \( \sigma_k \approx \beta \), where \( \beta \) is the 1/e divergence angle.](image-url)
In the present context we are more interested in the focal region of the beam, in the so-called Rayleigh range: in that waist like region close to the focus of the beam, the wave fronts appear nearly flat, much like in a plane wave. Strictly speaking, the beam wave in the focal region is not a transverse radiation wave such that \( \hat{s} = \hat{e} \times \hat{b} \). It can’t be, since it is a superposition of waves with varying polarization vectors \( \hat{e} \). However, for the present needs it is sufficient to consider only the so-called paraxial beams such that \( \sigma_k^2 \ll 1 \). This allows to neglect the \( z \)-components of the partial polarization vectors and to approximate the \( z \)-components of the vector \( \hat{k} \) of the partial waves as \( \kappa_z = \sqrt{1 - (\kappa_x^2 + \kappa_y^2)} \approx 1 - (\kappa_x^2 + \kappa_y^2)/2 \). Furthermore, for the present needs it is sufficient to allow for two restrictions: i) We send all partial waves through the same polarization filter. Because of the paraxial condition, we always get the same \( \hat{e} \) that does not depend on \( \kappa_x, \kappa_y \). ii) We restrict ourselves to the Rayleigh range (viz. Fig. 2.2), whose length is characterized by the so-called Rayleigh-parameter \( z_r = 1/k\sigma_k^2 \). Within the Rayleigh range, we can approximate the beam wave with a quasi-plane wave:

\[
\mathbf{E}(x, y, z) = \tilde{B}(kx, ky) E\hat{e} e^{ikz},
\]

where \( \tilde{B}(kx, ky) \) is a dimensionless spatial profile:

\[
\tilde{B}(kx, ky) = \frac{1}{2\pi} \int \int B(\kappa_x, \kappa_y) e^{ik_0 x + ik_0 y} d\kappa_x d\kappa_y.
\]

According to Parseval’s theorem, the distribution \( \tilde{B}(kx, ky) \) is quadratically normalized such that \( \int \int |\tilde{B}(kx, ky)|^2 k^2 dx dy = 1 \). Since the \( k \)-vector distribution \( B(\kappa_x, \kappa_y) \) and the transverse spatial profile \( \tilde{B}(kx, ky) \) are a pair of Fourier transforms, we expect the following general behavior: the larger is the \( k \)-vector variance \( \sigma_k^2 \), the smaller is the beam cross-section \( a_w^2 = \int \int (x^2 + y^2)|\tilde{B}(x, y)|^2 k^2 dx dy \). For the special case of Gaussian \( k \)-distribution from Eq. 2.31 one obtains

\[
\tilde{B}(kx, ky) = \frac{1}{\sqrt{\pi k^2 a_w^2}} e^{-(x^2 + y^2)/2a_w^2} \quad \text{where} \quad a_w^2 = \frac{1}{k^2 \sigma_k^2} = \frac{z_r}{k} \quad (2.37)
\]

In general one must expect \( k a_w \sigma_k \leq 1 \); the Gaussian beam, for which the equality applies, is a minimum uncertainty beam and therefore so special.

In subsequent applications we shall find it convenient to express the electric field of the beam wave in terms of a measurable quantity, namely the beam power \( P \) [W]. Combining Eq. 2.33 with Eq. 2.35 and setting \( \langle \kappa_z \rangle \approx 1 - \sigma_k^2/2 \approx 1 \), we write the electric field of the beam wave as

\[
\mathbf{E}(x, y, z) = \sqrt{\bar{P}} \sqrt{\frac{2}{\epsilon c_o}} X(x, y) \hat{e} e^{ikz}, \quad (2.38)
\]

where \( X(x, y) = k\tilde{B}(x, y) \). A convenient model for the profile \( X(x, y) \) is the Gaussian

\[
X(x, y) = \exp(-[x^2 + y^2]/2a_w^2)/\sqrt{\pi a_w^2}. \quad (2.39)
\]

The quantity \( \pi a_w^2 \) defines the effective cross-section of the beam. Thus Eq. 2.38 can be understood as

\[
\mathbf{E}(x, y, z) = \sqrt{\frac{2}{\epsilon c_o}} \sqrt{\bar{I}(x, y)} \hat{e} e^{ikz} = \sqrt{\frac{2}{\epsilon c_o}} \sqrt{\bar{I}X(x, y)} \hat{e} e^{ikz}, \quad (2.40)
\]

Here \( \bar{I}(x, y) = P|X(x, y)|^2 \) is the irradiance [W/m²]. On the right hand side we factored \( \bar{I}(x, y) \) into the mean irradiance \( \bar{I} = P/\pi a_w^2 \) and the dimensionless profile
\[ Y(x, y) = \exp(-[x^2 + y^2]/a_w^2), \] which characterizes the intensity distribution through the beam cross-section. Recall, however, that sooner or later the quasi plane wave will transform into the quasi spherical wave, Eq. 2.34, which radiates into a solid angle \( \Omega_w = 4\pi/k^2a_w^2 \). The radiant intensity \( I_\Omega [\text{W/sr}] \) of the beam is \( I_\Omega = P/\Omega_w \). Note that for a Gaussian beam \( I_\Omega = I z^2/4 \), which is yet another manifestation of the uncertainty relation.

It is obvious that the Rayleigh range is a good place in a beam to measure and manipulate the polarization. One only has to make sure that \( z_r \) is sufficiently large as compared with the thickness of the polarizing elements. Recall that for a paraxial beam \( z_r = ka_w^2 \approx 1/k\beta^2 \), where \( \beta \) is the 1/e divergence angle of the beam. In a typical laser beam with \( \beta \) of a couple of millirads one has \( z_r \) many meters long. On the other hand, when the beam is focused with a microscope objective, \( z_r \) may shrink to only a fraction of wavelength. The question arises, what happens with the polarization upon focusing the beam. Fortunately, the answer is favorable [11]: linear or circular polarization prepared in the Rayleigh range remains largely unaffected by focusing through a good lens (free of birefringence), up to numerical aperture close to 1.

### 2.3.3 Light sources, dipole fields

There is still a problem with the beams, as defined mathematically in Eq. 2.29: they do not yet represent physical reality, because they propagate from nowhere to nowhere. They do not have a source, where the energy is pumped into the light field, and they do not have a sink, where the radiation energy is absorbed and re-stored in the constituents of matter. The basic constituents of condensed matter are atoms and molecules whose sizes are much smaller than the light wavelength. The only way the electromagnetic field can interact with molecules is to deform their charge systems into dipoles. Such dipolar interactions are involved in both, scattering and absorption of light by molecules, but the effect of absorption on light propagation is rather trivial (a photon path is terminated). We are mainly interested in scattering: the impinging oscillating electric field induces in a target molecule an oscillating dipole, which in turns re-radiates the electromagnetic energy in form of a spherical wave. As usual, we assume that the field-molecule interaction is linear and that the impinging field is harmonic; all quantities oscillate with \( \exp(-i\omega t) \). We set aside the rare case of chiral molecules and consider only the induced electric dipole:

\[ d = A_m \cdot E^i \]  \hspace{1cm} (2.41)

This is the molecular version of the constitutive relation in homogeneous continuous media, Eq. 2.17. The polarizability tensor \( A_m \) represents the susceptibility of the molecule to the applied electric field. For convenience we also define a normalized version of \( A_m \), namely \( A = A_m/\alpha \), where \( \alpha = \text{Trace}(A_m)/3 \) is the mean polarizability. With isotropic molecules, the tensor \( A \) reduces to the unit matrix \( I \).

The electromagnetic field generated by a dipole source is given by the Hertz-solution of Maxwell equations, which represents the fundamental mode in the space of polarized spherical waves. For convenience we reproduce here the electric part of the scattered field. The scattering molecule is positioned in the origin of a coordinate system at \( r_m = 0 \). The scattered field can be written as the sum of two terms: \( E^s(r) = E^s_{\text{near}}(r) + E^s_{\text{rad}}(r) \). The first term represents the near field:

\[ E^s_{\text{near}}(r) = \frac{1}{4\pi\epsilon_0} k^3 e^{ikr} \left\{ \frac{1}{k^2 r^2} - i \frac{1}{kr} \right\} [3(d\cdot\hat{r})\hat{r} - d] \]  \hspace{1cm} (2.42)
Note that the near field contributions decay rapidly with $1/(kr)^2$ and $1/(kr)^3$. Near fields do not represent propagating photons. However, they mediate energy exchange at close distances between the molecules and therefore play an important role in optical properties of condensed matter. The second term is the radiation field:

$$E_{rad}^s(r) = \frac{1}{4\pi\epsilon_0} k^3 e^{i k r} \left[ d - \hat{r}(\hat{r} \cdot d) \right]$$

(2.43)

The radiation field $E_{rad}^s(r, t)$ is a polarized spherical wave $\sim \exp(-i \omega t + i k r)/k r$ whose polarization is given by the vector $[d - \hat{r}(\hat{r} \cdot d)] = d_\perp$. Note that this $d_\perp$ is the component of the induced dipole moment that is perpendicular to the director $\hat{r}$. Thus, the radiation field is a transverse field, just like a plane wave. Only now we have a spherical wave that propagates radially in the direction $\hat{s} = \hat{r}$. The oscillating dipole acts as a source of photons. The radiation power, i.e. electromagnetic energy emitted in unit time, is

$$P_d = \frac{1}{4\pi\epsilon_0} \frac{c}{3} k^4 |d|^2$$

(2.44)

Note here the $k^4$-dependence, characteristic for dipole radiation. Note also, that with the molecular dipole we introduced a fundamental inhomogeneity into the otherwise homogeneous medium.
Chapter 3

Manipulating polarization

“The type of optical system with which this chapter deals is not the more familiar type involving lenses, prisms, etc., but is rather the type which is composed of retardation plates, partial polarizers, and plates possessing the ability to rotate the plane of polarization. We shall therefore be concerned not with the directions of rays of light, but with the state of polarization and the intensity of the light as it passes through the optical system.”[13]

3.1 Linear and circular polarization

In chapter 2.2 we made an important observation: plane waves and quasi-plane waves are transversal waves: the field vector $\mathbf{E}(r,t) = E\hat{e}\exp(ik\cdot r - i\omega t)$ oscillates in a plane that is perpendicular to the propagation direction $\hat{s}$. One can always rotate the coordinate system so that $\hat{z}$-axis is oriented along the propagation direction $\hat{s}$ and therefore $\hat{e} = e_x\hat{x} + e_y\hat{y} + 0\cdot \hat{z} = e_x\hat{x} + e_y\hat{y}$. In other words, the polarization vector $\hat{e}$ can be expressed as a complex linear combination of the two orthogonal unit vectors $\hat{x}$ and $\hat{y}$. These two real unit vectors form the so-called linearly polarized base. Any linear combination with real coefficients, such as for example $\hat{e} = \cos(\psi)\hat{x} + \sin(\psi)\hat{y}$, represents linearly polarized light. The classical observable quantity is $\mathbb{R}[\hat{e}\exp(ik_s z - i\omega t)] = \hat{e}\cos(k_s z - \omega t) = [\cos(\psi)\hat{x} + \sin(\psi)\hat{y}]\cos(k_s z - \omega t)$. The real field vector oscillates linearly in a plane rotated by $\psi$ with respect to the $xz$-plane, as illustrated in Fig. 3.1 A. Combining $\hat{x}$ and $\hat{y}$ with complex coefficients, one obtains the general case of elliptically polarized light. Two particularly important combinations are $\hat{e}_L = \hat{x}/\sqrt{2} + i\hat{y}\sqrt{2}$ and $\hat{e}_R = \hat{x}/\sqrt{2} - i\hat{y}\sqrt{2}$. Note that $i = \exp(i\pi/2)$; there is a phase shift $\pm\pi/2$ between the two linearly polarized components:

$$\mathbb{R}[\hat{e}_{L/R}\exp(ik_s z - i\omega t)] = \frac{1}{\sqrt{2}} \cos(\omega t - k_s z)\hat{x} \pm \frac{1}{\sqrt{2}} \sin(\omega t - k_s z)\hat{y},$$

(3.1)

where again + is for $\hat{e}_L$ and - for $\hat{e}_R$. The geometry of the right handed polarization $\hat{e}_R$ is illustrated in Fig. 3.1 B. When frozen in time, the tips of the real part of the polarization vector along a ray form a right handed (continental) screw. The right part of Fig. 3.1 B illustrates the time rotation of the tip, as viewed in the direction of propagation: the vector rotates counter-clockwise, which is rather confusing. Circularly polarized states are fundamental in the photon picture of light: a photon in the left-handed state $\hat{e}_L \equiv |\sigma^+\rangle$ has its spin oriented in the direction of propagation $\hat{s}$, whereas $\hat{e}_R \equiv |\sigma^-\rangle$ is a photon state with negative spin. These two complex vectors $\hat{e}_L$ and $\hat{e}_R$ are orthogonal. They form the so-called circularly polarized base, but this choice of
base is slightly too abstract for the needs of an experimentalist. We shall keep working with the linear base \( \hat{x} \) and \( \hat{y} \).

3.2 Basic hardware elements

3.2.1 Retarder plate

The basic optical element for manipulating polarization is a retarder plate, a transparent plan-parallel plate made from birefringent material. The anisotropy axis \( \hat{a} \) is oriented in parallel with the faces of the plate and the plate is oriented normally to the \( \hat{k} \) of the impinging wave so that \( \hat{k} \perp \hat{a} \). Consequently the ordinary and the extraordinary waves propagate in the same direction \( \hat{s} \parallel \hat{k} \) but with different phase velocities \( c/n_{ord} \) and \( c/n_{ext} \). Upon traversing the plate the two waves will accumulate a mutual phase shift, which may turn linear polarization into elliptical, or vice versa. Fig. 3.2 explains the function of the retarder in little more detail and defines the coordinate system used in the subsequent formal analysis. In the complex vector notation the action of the retarder can be expressed by the following equation:

\[
\hat{e}(z_{out}) = e^{ik_xd} \hat{x} \left[ \hat{x} \cdot \hat{e}(z_{in}) \right] + e^{ik_yd} \hat{y} \left[ \hat{y} \cdot \hat{e}(z_{in}) \right]
\]  

(3.2)

Note here the two projection operators \( \hat{x} \hat{x} \cdot \) and \( \hat{y} \hat{y} \cdot \) that project the input state \( \hat{e}(z_{in}) \) on the states \( \hat{x} \) and \( \hat{y} \), respectively. The phasors \( \exp(ik_xd) \) and \( \exp(ik_yd) \) account then for the propagation over a distance \( d = z_{out} - z_{in} \) with two different phase velocities \( v_x \) and \( v_y \). There is no power lost in the plate and so the state \( \hat{e} \) remains a normalized unit vector after propagating through the plate. A particularly important special case of a retarder is the \( \lambda/4 \)-plate, whose thickness \( d \) is carefully adjusted so that \( (k_y - k_x)d = \)

![Figure 3.1](image)

Figure 3.1: A: Linearly polarized wave. Plotted is the real part of the polarization vector as the function of \( z \) at \( t = 0 \): \( \Re[\hat{e}\exp(ik_xz)] = [\cos(\psi)\hat{x} + \sin(\psi)\hat{y}] \cos(-k_zz) \). B: Right circularly polarized wave. The right handed spiral is the real part of the polarization vector: \( \Re[\hat{e}_R \exp(ik_xz)] = \frac{1}{\sqrt{2}} \cos(-k_zz) \hat{x} - \frac{1}{\sqrt{2}} \sin(-k_zz) \hat{y} \). As the spiral is pushed through the \( xy \)-plane (without rotation), the arrow \( \Re[\hat{e}_R \exp(-i\omega t)] = \frac{1}{\sqrt{2}} \cos(\omega t) \hat{x} - \frac{1}{\sqrt{2}} \sin(\omega t) \hat{y} \) rotates counter-clockwise. Note that the \( x \) and \( y \) axes appear reversed because we view the rotation in the direction of wave propagation.
Figure 3.2: Retarder plate. The anisotropy axis $\hat{a}$ is oriented in parallel with the faces of the plate, in our case along the $\hat{y}$-direction. Upon entering the plate, the linearly polarized input wave $E^{in}$ splits into the ordinary and extraordinary wave, but both beams propagate in the same direction $\hat{s}$, because $\hat{k}$ is normal to the interface and to $\hat{a}$ (Eqs. 2.22 and 2.25). The ordinary wave oscillates along $\hat{x}$ and propagates with phase velocity $c/n_x$, where $n_x = \sqrt{\varepsilon}$. The extraordinary wave oscillates along $\hat{y}$ and propagates with phase velocity $c/n_y$, where $n_y = \sqrt{\varepsilon + \Delta\chi}$. (In our example $\hat{a}$ is a slow axis, such as in a quartz retarder.) Upon passing the thickness $d$, the two waves accumulated a phase difference $\delta = d(k_y - k_x)$, so that they recombine into an elliptically polarized state, in our example nearly $\hat{e}_R$. Recall from electrodynamics textbooks that there are certain boundary conditions to be met at each interface. Because $\hat{k}$ is normal to the interface and to $\hat{a}$, only two boundary conditions are sufficient in our case: i) phase matching $\cos(k_{in}z_i) = \cos(k_xz_i + \phi_x) = \cos(k_yz_i + \phi_y)$, ii) tangential components of the $E$-vector are continuous: $E_{in} = E_x\hat{x} + E_y\hat{y}$.

$\pm \pi/2$. The special retarder equation reads

$$\hat{e}(z_{out}) = e^{ik_xd} \left\{ \hat{x} [\hat{x} \cdot \hat{e}(z_{in})] \pm i \hat{y} [\hat{y} \cdot \hat{e}(z_{in})] \right\}. \quad (3.3)$$

This looks much like circular polarization. Indeed, upon inserting a linearly polarized input state $\hat{e}(z_{in}) = \cos(\phi)\hat{x} + \sin(\phi)\hat{y}$ (and neglects the common phasor $\exp(ik_xd)$), one obtains $\hat{e}(z_{out}) = \cos(\phi)\hat{x} \pm i \sin(\phi)\hat{y}$. This is elliptical polarization. To obtain perfectly circular state, one must turn the anisotropy axis of the $\lambda/4$-plate in $45^\circ$ with respect to input polarization, so that $\cos(\phi) = 1/\sqrt{2}$ and $\sin(\phi) = \pm 1/\sqrt{2}$. Turning the plate means a rotation of the coordinate system. We define the rotated coordinate system $\hat{x}', \hat{y}'$ through

$$\hat{x}' = \cos(\phi)\hat{x} - \sin(\phi)\hat{y}' \quad (3.4)$$
$$\hat{y}' = \sin(\phi)\hat{x}' + \cos(\phi)\hat{y}'$$

The angle $\phi$ is measured clockwise when looking along the direction of propagation. In the rotated coordinate system the anisotropy axis $\hat{a}$ is still oriented along $\hat{y}'$ so that

$$\hat{e}'(z_{out}) = e^{ik_xd} \left\{ \hat{x}' [\hat{x}' \cdot \hat{e}'(z_{in})] \pm i \hat{y}' [\hat{y}' \cdot \hat{e}'(z_{in})] \right\}. \quad (3.5)$$

However, $\hat{e}'(z_{in})$ must be expressed in the new coordinate system:

$$\hat{e}'(z_{in}) = \cos(\psi) [\cos(\phi)\hat{x}' - \sin(\phi)\hat{y}'] + \sin(\psi) [\sin(\phi)\hat{x}' + \cos(\phi)\hat{y}']. \quad (3.6)$$

It is easy to verify that circular polarization is recovered when $\psi - \phi = \pm \pi/4$. Finally, one would have to transform $\hat{e}'(z_{in})$ back into the laboratory coordinate system, but vector algebra is a somewhat clumsy tool for such calculations.
3.2.2 Linear polarizer

A longer look at the retarder equation, Eq. 3.2, suggests how to make a polarization filter: if one could equip the plate with some device which blocks off one of the terms (say the extraordinary term, \( \hat{y} \)), then the device would simply project the incoming state \( \hat{e}(z_{in}) \) on the linear state \( \hat{x} \):

\[
e(z_{out}) = e^{ik_x d} \hat{x} \left[ \hat{x} \cdot \hat{e}(z_{in}) \right] \quad (3.7)
\]

In other words, no matter what \( \hat{e}(z_{in}) \) is, the output would always be the linearly polarized state \( \hat{x} \) multiplied with the complex amplitude \( \left[ \hat{x} \cdot \hat{e}(z_{in}) \right] = e^x \). Of course the output of a filter is no longer a unit vector, since \( e^*(z_{out}) \cdot e(z_{out}) = |e^x(z_{in})|^2 \neq 1 \); we lost some of the beam power in the filter.

One way to produce a linear polarizer is Glan Thompson prism shown in Fig. 3.3: the retarder plate (made from calcite) is cut into two triangular prisms and glued again together with a cement whose refractive index matches the smaller of \( n_{ext} = 1.486 \) and \( n_{ord} = 1.658 \). Thus the extraordinary ray is perfectly transmitted. The ordinary ray, however, undergoes total internal reflexion on the calcite-cement-interface, if the cut angle is appropriate. Another way is to make either \( k_x \) or \( k_y \) in Eq. 3.2 imaginary (say \( k_x \)), so that the amplitude of the chosen component decays exponentially with \( \exp(-|k_x|d) \). A well known example of this principle is the Polaroid sheet made of strongly anisotropic absorbing material. Because of the strong anisotropy, the material exhibits strong linear dichroism, i.e. the extraordinary wave is much more absorbed than the ordinary wave. Modern sheet polarizers are practically as perfect as the Glan-Thompson prism. However, because of the strong absorption, they can’t be used in a laser beam. On the other hand, the sheet form is perfect for polarization imaging.

By combining linear polarizers with retarder plates, one achieves the full control of the polarization state. In particular, one can produce and calculate filters for circularly polarized states. However, as already mentioned, the vector notation is not well suited for such calculations. They can be best done using the elegant tools to be discussed in the next chapter.

3.3 Jones formalism

In 1940, R. Clark Jones, researcher with Polaroid Corporation and Harvard scholar, got annoyed with vector calculus, which prompted him to design an efficient formalism for the analysis of light propagation through optical elements such as, for example, sheets of Polaroid foil [13]. Jones represents optical elements as matrix operators \( O \), which act on the polarization states that are represented by 2D complex column vectors \( J \) (one column matrices, known today as the Jones vectors): \( J_{out} = O \cdot J_{in} \). There is no doubt that Jones was inspired by the matrix formulation of quantum mechanics. Thus,
we shall introduce the Jones formalism using Dirac’s bra-ket-notation, \( \langle \text{bra} | \text{and} | \text{ket} \rangle \). For now, this is useful in order to distinguish the Jones vectors from 3D field vectors such as \( \mathbf{E} \). Later we shall fill the bra-kets with a more general content, without having to change the notation. Using the linearly polarized base, the Jones vector \( |E\rangle \) that corresponds to the field vector \( \mathbf{E} \) reads

\[
\mathbf{E} \equiv |E\rangle = \begin{pmatrix} E_x \\ E_y \end{pmatrix}
\] (3.8)

Here \( E_x \) and \( E_y \) are the complex components of the \( \mathbf{E} \)-vector in a local coordinate system whose \( \hat{z} \)-axis coincides with the propagation direction \( \hat{s} \). Each ket has a companion bra, a complex conjugate and transpose of the ket:

\[
|E^*\rangle \equiv \langle E | = \begin{pmatrix} E_x^* \\ E_y^* \end{pmatrix}
\] (3.9)

In the bra-ket-notation, the scalar product of two complex vectors \( |a\rangle \) and \( |b\rangle \) reads

\[
\mathbf{a}^* \cdot \mathbf{b} \equiv \langle a | b \rangle = \begin{pmatrix} a_x^* \\ a_y^* \end{pmatrix} \begin{pmatrix} b_x \\ b_y \end{pmatrix} = a_x^* b_x + a_y^* b_y
\] (3.10)

For now, \( \langle a | b \rangle \) is merely an abbreviation for \( a_x^* b_x + a_y^* b_y \), but later we shall employ a more general definition of the scalar product. A special case of a scalar product is the norm \( \langle a | a \rangle \equiv | \mathbf{a} |^2 \). A unit Jones vector \( |e\rangle \) is normalized such that \( \langle e | e \rangle = 1 \), just like an ordinary 3D unit vector in Euclidean space. Recall that unit vectors are dimensionless. For now, a complex unit Jones vector \( |e\rangle \) represents a polarization state, equivalent to a polarization unit vector \( \hat{e} \): \( \mathbf{E} = \mathbf{E} \hat{e} \equiv |E\rangle = E |e\rangle \). The following three pairs of polarization states are particularly important:

\[
|X\rangle := \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad |Y\rangle := \begin{pmatrix} 0 \\ 1 \end{pmatrix} \\
|\rangle \rangle := \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix} \quad |\ \rangle := \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix} \\
|R\rangle \equiv |\sigma^{-}\rangle := \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -i \end{pmatrix} \quad |L\rangle \equiv |\sigma^{+}\rangle := \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ i \end{pmatrix}
\] (3.11)

The states \( |X\rangle \), \( |Y\rangle \) and \( |\rangle \rangle \), \( |\ \rangle \) are linearly polarized; both pairs are orthogonal and form two linearly polarized bases: any Jones vector \( |E\rangle \) can be expressed as

\[
|E\rangle = E_x |X\rangle + E_y |Y\rangle,
\] (3.12)

where \( E_x \) and \( E_y \) are complex scalars. The states \( |\rangle \rangle \), \( |\ \rangle \) are easily generated by turning the polarizer \( \pm 45^\circ \) with respect to \( x \)- and \( y \)-axes. At the first glance it appears superfluous to have two linear bases, but the significance of the turned base will emerge later. The states \( |L\rangle \), \( |R\rangle \) form the circularly polarized base.

### 3.4 Optical elements in Jones language

An optical element is a Jones matrix \( \mathbf{O} \) that acts on an input state \( |E^{\text{in}}\rangle \) and produces an output state \( |E^{\text{out}}\rangle \) according to \( |E^{\text{out}}\rangle = \mathbf{O} |E^{\text{in}}\rangle \). The general structure of a Jones matrix in the \( x, y \)-base is

\[
\mathbf{O} = \begin{pmatrix} O_{xx} & O_{xy} \\ O_{yx} & O_{yy} \end{pmatrix}
\] (3.13)
The numbering of the matrix elements is such that, e.g., $O_{xy}$ means the contribution of the $y$-component of the input state to the $x$-component of the output state:

$$
E_{ox} = O_{xx} E_{in}^x + O_{xy} E_{in}^y
$$

$$
E_{oy} = O_{yx} E_{in}^x + O_{yy} E_{in}^y
$$

(3.14)

### 3.4.1 Projecting filters

The central concept in the present approach is the projection operator. A projector is constructed by multiplying a column vector from the right by a row vector, which results in a hermitian 2x2 matrix:

$$
\hat{f} \hat{f}^\dagger \equiv |f\rangle \langle f| := \begin{pmatrix} f_x & f_y \\ f_y^* & f_x^* \end{pmatrix}
$$

(3.15)

Such matrices represent polarization filters. Note that a filter state is normalized such that $\langle f| f\rangle = 1$. In optics the projection operator is known under the name “Jones matrix of the filter”. Three particularly important pairs of filter matrices are readily constructed from Eqs. 3.11:

$$
F_X = |X\rangle \langle X| = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}
$$

$$
F_Y = |Y\rangle \langle Y| = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}
$$

(3.16)

$$
F_L = |L\rangle \langle L| = \begin{pmatrix} \frac{1}{\sqrt{2}} & i \frac{1}{\sqrt{2}} \\ -i \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix}
$$

$$
F_R = |R\rangle \langle R| = \begin{pmatrix} 1 & 0 \\ 0 & -i \end{pmatrix}
$$

(3.17)

$$
F_{/} = 1 \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}
$$

$$
F_{\backslash} = 1 \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}
$$

(3.18)

$F_X$, $F_Y$, $F_{/}$ and $F_{\backslash}$ represent linear polarizers. Thanks to Polaroid, a linear polarizer is a common device, but circular polarizers $F_L$ and $F_R$ require some construction job.

### 3.4.2 Transformers

To transform the polarization from one state to the other, one needs two construction elements: i) Retarder plate of variable thickness (either fast or slow anisotropy axis will do), to generate a phase shift between the ordinary and extraordinary beams, ii) Rotational stage, to transfer the optical power between the two orthogonal states. The Jones matrix of a retarder is readily obtained by translating Eq. 3.2 into the Jones language. The result is a linear combination of two projectors:

$$
|e(z_{out})\rangle = |e^{ik_x d} F_x + e^{ik_y d} F_y |e(z_{in})\rangle = T_{\text{Ret}} |e(z_{in})\rangle
$$

(3.19)

Omitting a common phasor $\exp(ik_x d)$, the retarder matrix reads:

$$
T_{\text{Ret}} = \begin{pmatrix} 1 & 0 \\ 0 & \exp(i\Delta k d) \end{pmatrix}
$$

where $\Delta k = k_y - k_x$

(3.20)

Two kinds of retarders are readily available: $\lambda/4$-plates, whose thickness is accurately adjusted so that $\Delta k d = \pm \pi/2$, and $\lambda/2$-plates, whose thickness is such that $\Delta k d = \pm \pi$. The sign is + for a slow anisotropy axis ($\Delta \chi > 0$). Assuming that $\hat{a} \parallel \hat{y}$, Jones matrices of these two retarders read:

$$
T_{\lambda/4} = \begin{pmatrix} 1 & 0 \\ 0 & \pm i \end{pmatrix}
$$

$$
T_{\lambda/2} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}
$$

(3.21)
The $\lambda/4$-plates are used for transforming the polarization from linear to circular (and vice versa), and $\lambda/2$-plates are needed to rotate the plane of linear polarization. The last element we need is a tool to rotate the retarder plates or linear polarizers:

$$ R_2(\phi) = \begin{pmatrix} \cos(\phi) & +\sin(\phi) \\ -\sin(\phi) & \cos(\phi) \end{pmatrix} \quad R_2^{-1}(\phi) = \begin{pmatrix} \cos(\phi) & -\sin(\phi) \\ +\sin(\phi) & \cos(\phi) \end{pmatrix} \quad (3.22) $$

The angle $\phi$ is measured clockwise, when viewing in the direction of propagation. For future algorithmic needs we also recall that the rotations can be expressed in terms of the direction cosines:

$$ R_2(\phi) = \begin{pmatrix} \hat{x}'\cdot\hat{x} & \hat{x}'\cdot\hat{y} \\ \hat{y}'\cdot\hat{x} & \hat{y}'\cdot\hat{y} \end{pmatrix} \quad R_2^{-1}(\phi) = \begin{pmatrix} \hat{x}\cdot\hat{x}' & \hat{x}\cdot\hat{y}' \\ \hat{y}\cdot\hat{x}' & \hat{y}\cdot\hat{y}' \end{pmatrix} \quad (3.23) $$

When applied to a Jones vector whose components are given in laboratory coordinates $\hat{x}$, $\hat{y}$, the operation $|J\rangle = R_2 |J\rangle$ expresses this vector in a local system $\hat{x}'$, $\hat{y}'$ that is rotated by $\phi$ around the $z$-axis (rotation clockwise when viewing in the direction of propagation). The inverse matrix $R_2^{-1}$ transforms the components that are known in the rotated system into the components in the laboratory system: $|J\rangle = R_2^{-1} |J'\rangle$. The subscript $\phi$ is a reminder of the fact that the rotations are 2D: all components of the optical system under consideration are well aligned on the $\hat{z}$-axis and the anisotropy axis $\hat{a}$ is normal to $\hat{z}$. The operation of rotating an optical element is mathematically expressed as

$$ O(\phi) = R_2^{-1}(\phi)O(0)R_2(\phi). \quad (3.24) $$

From right to left the first rotation brings us into the $\hat{x}'$, $\hat{y}'$-system, the inverse rotation rotates the output back into the laboratory system. In the same way as a linear retarder, one can also construct a circular retarder made from a chiral medium, Sec. 2.2.2. The only difference is that the two filters are $F_L$ and $F_R$ (instead of $F_X$ and $F_Y$). Introducing the abbreviation $\delta k = k_L - k_R$, we obtain the chiral transfer matrix $T_{Chi} = e^{ikLd}F_L + e^{ikRd}F_R$ as

$$ T_{Chi} = e^{ikd} \begin{pmatrix} \cos(\delta k d) & +\sin(\delta k d) \\ -\sin(\delta k d) & \cos(\delta k d) \end{pmatrix}. \quad (3.25) $$

This matrix represents the rotation of linear polarization, i.e., “optical activity”.

### 3.4.3 Producing circular polarization

In order to produce circularly polarized light out of linearly polarized laser light we need two optical elements: i) an optional $\lambda/2$-retarder in a rotary stage, which allows to align the polarization in $\hat{x}$-direction, if necessary. ii) A $\lambda/4$-retarder whose $\hat{a}$-axis is aligned in $\pm \pi/4$ with $\hat{x}$. In this way the linear polarization is splitted symmetrically into the ordinary and extraordinary modes, arranged in a V-shaped manner around $\hat{x}$. The $\pi/2$-phase difference upon passing the $\lambda/4$-plate produces circularly polarized output. In Jones formalism this operations are expressed as $|e(z_{out})\rangle = T_{\lambda/4}(\pi/4)T_{\lambda/2}(\phi) |e(z_{in})\rangle$. The action of the rotating $\lambda/2$-plate $T_{\lambda/2}(\phi)$ on an arbitrary linearly polarized state $\langle e(z_{in})| = (\cos(\psi) \sin(\psi))$ is

$$ |e(z_{out})\rangle = T_{\lambda/2}(\phi) |e(z_{in})\rangle = R_2^{-1}(\phi)T_{\lambda/2}R_2(\phi) |e(z_{in})\rangle \quad (3.26) $$

$$ \begin{pmatrix} \cos(2\phi - \psi) \\ -\sin(2\phi - \psi) \end{pmatrix} = \begin{pmatrix} \cos(2\phi) & +\sin(2\phi) \\ -\sin(2\phi) & -\cos(2\phi) \end{pmatrix} \begin{pmatrix} \cos(\psi) \\ \sin(\psi) \end{pmatrix} \quad (3.27) $$
Now we align the polarization by setting \(2\phi = \psi\) and investigate the action of the \(\lambda/4\)-plate \(T_{\lambda/4}(\pi/4)\) on the input state \(\langle e(z_{\text{in}})\rangle = (1\ 0)\):

\[
\langle e(z_{\text{out}})\rangle = T_{\lambda/4}(\pi/4) \langle e(z_{\text{in}})\rangle = R^{-1}_2(\pi/4) T_{\lambda/4} R_2(\pi/4) \langle e(z_{\text{in}})\rangle
\]

\[
e^{i\pi/4} |R\rangle = e^{i\pi/4} \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -i \end{pmatrix} = e^{i\pi/4} \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -i \\ 1 \\ 0 \end{pmatrix}
\]

We succeeded in producing right-circularly polarized light! The phasor \(\exp(i\pi/4)\) is usually irrelevant.

### 3.4.4 Blocking and filtering circular polarization

In diagnostic applications we are interested in the role of circular polarization in the propagation and interaction with tissues. Thus, we must equip the detector with a device which is capable of transmitting selectively one circular polarization state and blocking the other. Having understood the role of the \(\lambda/4\)-plate as a polarization transformer, the construction is easy: we place into the beam line first a \(\lambda/4\)-plate \(T_{\lambda/4}(\pi/4)\) and then a linear polarization filter \(F_{X/Y}\); the order is in the direction of light propagation. We take the filter \(F_Y\) and investigate how the combined device acts on the circularly polarized states \(\langle R\rangle = \sqrt{1/2} (1 - i)\) and \(\langle L\rangle = \sqrt{1/2} (1 i)\). The short-hand formula is

\[
\langle e(z_{\text{out}})\rangle = F_Y T_{\lambda/4}(\pi/4) \langle e(z_{\text{in}})\rangle
\]

Inserting here some results from above we get

\[
F_Y T_{\lambda/4}(\pi/4) = e^{i\pi/4} \frac{1}{\sqrt{2}} \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ -i \\ 1 \\ 0 \end{pmatrix} = e^{i\pi/4} \frac{1}{\sqrt{2}} \begin{pmatrix} 0 & 0 \\ -i & 1 \end{pmatrix}
\]

Thus the output states for right and left circular polarization are

\[
\langle e(z_{\text{out}})\rangle = e^{i\pi/4} \frac{1}{2} \begin{pmatrix} 0 \\ -i \\ 1 \\ 1 \end{pmatrix} = i e^{i\pi/4} \begin{pmatrix} 0 \\ 1 \end{pmatrix}
\]

\[
\langle e(z_{\text{out}})\rangle = e^{i\pi/4} \frac{1}{2} \begin{pmatrix} 0 \\ -i \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}
\]

While left circular polarization is completely blocked off, right circular polarization is completely transmitted!! Note, however, that the device is not a projecting filter from Eq. 3.18, i.e. not a circular polarizer: the output state is not \(|R\rangle\), but rather the linearly polarized state \(|Y\rangle\). To make the device into a circular polarizer, we would have to add an appropriately oriented \(T_{\lambda/4}\). Note, however, that the finally measured observable is the power transmitted through the filter and not the polarization vector itself.

### 3.5 Deflectors

Jones’ original idea was aimed at the analysis of one dimensional optical systems consisting of polarizing elements that are neatly aligned on a single optical axis. Soon, however, the need came up to extend the concept to include elements that change the direction of light propagation, such as deflecting mirrors, refracting prisms, beam splitters or scattering particles. Since the deflected light still remains a transversal wave (in any case in the far field), we can still employ Jones formalism. One can define a Jones matrix \(T\) for any arbitrary two pairs of base vectors, as indicated on the left
hand side of Fig. 3.4. The element $T_{x'y'}$, for example, means then the contribution of the $y'$-component of the input state into the $x'$-component of the output state:

$$|E_{\text{out}}^{\text{out}}\rangle = \begin{pmatrix} T_{x'x'} & T_{x'y'} \\ T_{y'x'} & T_{y'y'} \end{pmatrix} |E_{\text{in}}^{\text{in}}\rangle$$

or explicitly

$$E_{\text{out}}^{x''} = T_{x'x'} E_{\text{in}}^{x'} + T_{x'y'} E_{\text{in}}^{y'}$$

$$E_{\text{out}}^{y''} = T_{y'x'} E_{\text{in}}^{x'} + T_{y'y'} E_{\text{in}}^{y'}$$

(3.34)

However, a wide spread convention is to use a special choice of the base vectors, which is indicated on the right hand side of Fig. 3.4: one chooses a common $\hat{x}$-axis that is perpendicular to the plane $\pi$ spanned by the propagation vectors $\hat{s}_i$ and $\hat{s}_o$. This plane $\pi$ is called “scattering plane” or “plane of incidence”. This $\pi$-system is not always convenient. If there is some symmetry axis in lab frame, like the anisotropy axis $\hat{a}$ or an overall system axis $\hat{s}_L$, then one may prefer a reference plane $\alpha \equiv \text{span}(\hat{s}_i, \hat{a})$ or $\zeta \equiv \text{span}(\hat{s}_i, \hat{s}_L)$ [14, 15]. However, if both, some $\hat{a}$ and some $\hat{s}_L$, are present, then one better reverts to the $\pi$-system.

![Figure 3.4: Specifying linearly polarized basis for a deflector.](image)

A good understanding of the geometrical relationships in Fig. 3.4 is essential for polarized Monte Carlo simulations considered in Chapter 7: Deflecting optical elements, such as scatterers, act in two ways: they transform the polarization state, but they also change the direction of propagation by an angle $\theta$ so that $\cos(\theta) = \hat{s}_i \cdot \hat{s}_o$. Thus, a deflection event must be characterized by two transformation matrices: i) the 2D Jones matrix that operates on the transverse field amplitudes, ii) the 3D geometrical rotational matrix $R(\phi, \theta) = R(\theta)R(\phi)$. Here $R(\phi)$ is the rotation by $\phi$ around $\hat{s}_i$, which achieves the transformation from $\hat{x}_i', \hat{y}_i', \hat{s}_i$ into $\hat{x}_i, \hat{y}_i, \hat{s}_i$. The rotation $R(\theta)$ around $\hat{x}_i$ transforms then from the transverse coordinate system $i$ to the coordinate system $i + 1$ of the next straight path segment. The position and orientation of the elements is specified in the laboratory coordinate system $\hat{x}_L, \hat{y}_L, \hat{z}_L$, and therefore we must keep track of the overall transformation as it accumulated until $i$.

### 3.5.1 Snellius and Fresnel laws in Jones language

A familiar, yet not quite trivial example of a deflecting element is an interface between two dielectric media with refractive index mismatch, such as the air-tissue interface. The geometry of the problem is outlined in Fig. 3.5. An interface between two media with refractive indices $n_i$ (for incoming) and $n_t$ (transmitted) is illuminated by a laser beam with the propagation vector $\hat{s}_i$. The interface is optically flat on the length scale
that is larger than the diameter of the beam. A part of the beam is reflected specularly into the new direction \( \hat{s}_r \), and a part is transmitted and refracted according to Snellius in the direction \( \hat{s}_t \). The vectors \( \hat{s}_i \), \( \hat{s}_r \) and \( \hat{s}_t \) lay in the plane \( \pi \); recall Sec. 3.5. In \( \pi \) lays also the surface normal \( \hat{n} \). To orient this \( \hat{n} \) unambiguously, we shall employ the convention indicated in Fig. 3.5: \( \hat{n} \) points in the direction of the gradient of refractive index. The common axis \( \hat{x} \) is then defined through \( x = \hat{s}_i \times \hat{n} / |\hat{s}_i \times \hat{n}| \). Note in Fig. 3.5 the surface tangent \( \hat{t} \), such that \( \hat{x} \), \( \hat{t} \) and \( \hat{n} \) span a right handed coordinate system. In the coordinate system \( \hat{x}, \hat{t}, \hat{n} \) the three propagation vectors can be expressed as

\[
\begin{align*}
\hat{s}_i &= (0, s_{ti}, s_{ni}) = (0, \sin(\theta_i), \cos(\theta_i)) \\
\hat{s}_r &= (0, s_{ti}, -s_{ni}) = (0, -\sin(\theta_i), \cos(\theta_i)) \\
\hat{s}_t &= \left(0, \frac{n_i}{n_t}s_{ti}, \sqrt{1 - \frac{n_i^2}{n_t^2}s_{ti}^2}\right) = \frac{n_i}{n_t}\left(0, \sin(\theta_i), \sqrt{\frac{n_t^2}{n_i^2} - \sin^2(\theta_i)}\right)
\end{align*}
\]

where \( s_{ti} = \hat{s}_i \cdot \hat{t} \) and \( s_{ni} = \hat{s}_i \cdot \hat{n} \). The expression for the transmitted propagation vector \( \hat{s}_t \) includes the Snellius law. Exploiting Eqs. 3.35 we can immediately write the cosines of the deflection angles \( \cos(\theta_r) = \hat{s}_i \cdot \hat{s}_r \) and \( \cos(\theta_t) = \hat{s}_i \cdot \hat{s}_t \). The angles \( \theta_t, \theta_r \), together with \( \phi \) specify completely the 3D transformation matrix, recall Sec. 3.5. We assume now that the rotation \( R(\phi) \) into the \( \pi \)-plane has been already done.

The next step is to obtain the Jones matrix that operates on the polarization. Recall that the field components \( E_\parallel \equiv E_y \) and \( E_\perp \equiv E_x \), whose polarization are parallel and perpendicular to the plane \( \pi \), exhibit different reflection and transmission coefficients. The coefficients, found in any textbook, can be combined in the Jones matrices for reflection and transmission. Using the abbreviations \( C_t = (\hat{s}_t \cdot \hat{n}) = \cos(\theta_t) \) and \( C_r = (\hat{s}_r \cdot \hat{n}) = \cos(\theta_r) = \sqrt{1 - \sin^2(\theta_t)}^2 \) we write

\[
\begin{align*}
T_r &= \begin{pmatrix} r_\perp & 0 \\ 0 & r_\parallel \end{pmatrix} \quad \text{where} \quad r_\perp &= \frac{n_t}{n_i} C_t C_r - n_i C_t, \\
&\quad \quad r_\parallel &= \frac{n_i}{n_t} C_t C_r + n_i C_t
\end{align*}
\]

\[
\begin{align*}
T_t &= \begin{pmatrix} t_\perp & 0 \\ 0 & t_\parallel \end{pmatrix} \quad \text{where} \quad t_\perp &= \frac{2n_t}{n_i} C_t C_r, \\
&\quad \quad t_\parallel &= \frac{n_i}{n_t} C_t C_r + n_i C_r
\end{align*}
\]

For \( \sin(\theta_t) \) we recall the Snellius law \( n_1 \sin(\theta_t) = n_2 \sin \theta_t \). A special case is total internal reflection, where \( \sin^2(\theta_t) > 1 \) and \( \cos(\theta_t) \) thus imaginary. In this case \( t_\perp = t_\parallel = 0 \)

---

Figure 3.5: Geometry of Fresnel’s laws at an interface.
(evanescent waves do not contribute to light propagation) and

\[ r_\perp = e^{i\delta_\perp} \quad \delta_\perp = 2 \arctan \left\{ \frac{n_t |C_t|}{n_i C_i} \right\} \]  

(3.38)

\[ r_\parallel = e^{i\delta_\parallel} \quad \delta_\parallel = 2 \arctan \left\{ \frac{n_i |C_i|}{n_t C_t} \right\} \]  

(3.39)

The corresponding transmittance and reflectance, i.e., transmission and reflection probabilities, are

\[ P_r = \frac{\langle E_r^r | E_r^r \rangle}{\langle E_i^i | E_i^i \rangle} = |r_\perp|^2 |e_{x_i}|^2 + |r_\parallel|^2 |e_{y_i}|^2 = |e_r|^2 \]  

(3.40)

\[ P_t = \frac{\langle E_t^r | E_t^r \rangle}{\langle E_i^i | E_i^i \rangle} = \frac{n_t C_t}{n_i C_i} \left\{ |t_\perp|^2 |e_{x_i}|^2 + |t_\parallel|^2 |e_{y_i}|^2 \right\} = \frac{n_t C_t}{n_i C_i} |e_t|^2, \]  

(3.41)

where \( e_{x_i} \) and \( e_{y_i} \) are the components of the complex unit polarization vector of the input, and \( e_r, e_t \) are the polarization vectors resulting thereof (not unit vectors). Just a little bit more work is needed when it comes to the algorithmic realization of an interface in the course of polarized MC simulations: The photon state is not known in the convenient \( \pi \)-base \( \hat{x}, \hat{y} \) but in some \( \hat{x}_i^t, \hat{y}_i^t \), which was the outcome of a previous scattering event. Thus, one must first determine \( \hat{x}, \hat{y} \).
Chapter 4

Measuring polarization

4.1 Basic concepts

The information contained in a Jones vector $|E\rangle$ consists of four numbers, namely the real and imaginary parts of the two complex components $E_x$ and $E_y$, or their magnitudes $|E_x|, |E_y|$ and their phases $\phi_x, \phi_y$. In optics the field components are not directly accessible, the measurables are proportional to the complex squares of the field components. However, most of the desired information can be extracted by appropriate filtering. It is important to realize that two kinds of filtering are actually involved: prior applying the polarization filters, one must first filter a transversal wave out of the impinging light field, see Sec. 5. Here we assume that this spatial filtering is already done. The light field whose polarization is to be measured is a collimated beam that carries the power $P$, as expressed in 2.33. The measured signal $S$ is proportional to the energy $U$ or the number of photons $N = U/\hbar\omega$, that cross the detector surface $A\hat{n}$ during a finite measuring time $T$. If the detector captures the entire beam, then

$$S \sim U = \int_A \int_T \mathbf{S} \cdot \mathbf{n} dA dT = \int_T |E(t)|^2 dt = T \langle P \rangle_t$$  \hspace{1cm} (4.1)

Here $\mathbf{S}$ is the Poynting vector. On the right hand side we recalled Eq. 2.33 and introduced also the time averaged power $\langle P \rangle_t$. Note that we are anticipating the possibility that the field vector $\mathbf{E}(t) = E(t)\hat{e}(t)$ varies during the measuring time. The measurement involves averaging over time! We shall soon return to this point, but for now we set $\mathbf{E}(t) = \mathbf{E} = \text{const}$, so that $S \sim |E|^2$. Often this $|E|^2$ is called intensity, denoted with $I$. “Intensity” is probably the most abused term in physics and therefore we use $S$ for the measured signal. However, in the present chapter we neglect all pre-factors, setting $S$ equal to the norm of a Jones vector $|E\rangle$:

$$S = \langle E|E \rangle = |E_x|^2 + |E_y|^2, \quad \text{where} \quad |E \rangle = \begin{pmatrix} E_x \\ E_y \end{pmatrix} = \sqrt{S} \begin{pmatrix} e_x \\ e_y \end{pmatrix}$$  \hspace{1cm} (4.2)

Upon equipping the detector with a polarization filter $|f\rangle$ the measured signal becomes

$$S_f = \langle E'|E' \rangle = \langle E|f \rangle \langle f|E \rangle = \langle E|f \rangle \langle f|E \rangle = S |\langle f|e \rangle|^2$$  \hspace{1cm} (4.3)

Here $|\langle f|e \rangle|^2$ is the probability that a photon with polarization $|e\rangle$ will pass through the filter $|f\rangle$. Note that we do not need the Jones matrix of the filter to calculate $S_f$, but only the scalar product of the state $|e\rangle$ with the filter $|f\rangle$. However, we need a filter set appropriate for the extracting of the full available information on the Jones vector
\[ |f \rangle = \left( \frac{\cos(\psi)e^{-i\beta/2}}{\sin(\psi)e^{+i\beta/2}} \right) \] (4.4)

The signal measured upon transmitting the state \( |E \rangle \) through this filter is

\[ S_f = \langle E^f | f^f \rangle = |E_x| \cos(\psi)^2 + |E_y| \sin(\psi)^2 + 2 \cos(\psi) \sin(\psi) \Re \{ E_x E_y^* e^{i\beta} \} \] (4.5)

Upon examining Eq. 4.5 we identify four particularly convenient special cases. Here they are compiled together with the corresponding hardware filter elements:

\[
\begin{align*}
\psi &= 0 : & |f \rangle &= |X \rangle & S_X &\sim |E_x|^2 \\
\psi &= \pi/2 : & |f \rangle &= |Y \rangle & S_Y &\sim |E_y|^2 \\
\psi &= \pi/4, \beta = 0 : & |f \rangle &= |/ \rangle & S_f &\sim \frac{1}{2} S_0 + \frac{1}{2} [E_x E_y^* + E_x^* E_y] \\
\psi &= \pi/4, \beta = \pi/2 : & |f \rangle &= |R \rangle & S_R &\sim \frac{1}{2} S_0 - \frac{i}{2} [E_x E_y^* - E_x^* E_y]
\end{align*}
\]

where \( S_0 = |E|^2 = |E_x|^2 + |E_y|^2 \). Obviously, the components \( |E_x|^2 \) and \( |E_y|^2 \) are extracted using the two orthogonal linear analyzers \( |X \rangle \) and \( |Y \rangle \). Note also that

\[
\begin{align*}
[E_x E_y^* + E_x^* E_y] &= 2 \Re E_x E_y^* = 2|E_x||E_y| \cos(\delta) \\
-i [E_x E_y^* - E_x^* E_y] &= 2 \Im E_x E_y^* = 2|E_x||E_y| \sin(\delta).
\end{align*}
\] (4.10) (4.11)

The angle \( \delta = \phi_x - \phi_y \) represents the phase difference between the two linear polarizations; \( \delta = n_{\text{even}} \pi/2 \) means linear polarization, \( \delta = n_{\text{odd}} \pi/2 \) circular. Note that two filters are needed to extract \( \delta \) including its sign. We take one of the circular polarization filters, e.g. \( |R \rangle \) (in practice a circular blocker will do). The second filter is one of the linear filters \(|/ \rangle \) \(|\rangle \). It is easy to solve the equations 4.6-4.9 for the quantities \( S_0, |e_x|, |e_y|, \) and \( \delta \), which fully characterize the polarized harmonic plane wave. Note that the choice of filters in Eqs. 4.6 is not unique, one could as well use \(|\rangle \) \(|\rangle \) and \(|L\rangle \).

### 4.2 Time averaging, Stokes parameters

At this point, we must clarify a subtle but important issue. So far we were concerned with perfectly coherent, pure photon states \( |E \rangle \) corresponding to harmonic waves. Such states are always perfectly polarized, i.e., they possess a unique polarization vector \( |e \rangle \). This may be a good approximation for a typical laser beam, but certainly not for an old-fashioned light bulb, which produces the so-called “natural unpolarized light”. Light reflected from a biological tissue, though not completely depolarized, is still quite different from the laser light. In modern terminology one would say that most of light sources generate state mixtures. In classical electrodynamics the mixture of states can be modeled by replacing the monochromatic wave \( E \exp(-i\omega t) \) with a quasimonochromatic wave \( E(t) \exp(-i\omega t) \), where the field vector \( E(t) \) fluctuates randomly on a time scale which is much longer than the oscillation period \( 1/\omega \) but much shorter than the experimental integration time \( T \). The measured quantities are then time averages, for example \( \langle S_f \rangle = \langle |\langle f | E \rangle|^2 \rangle \).

Even before the nature of light was fully understood, Sir George Gabriel Stokes realized this fact and suggested to organize the time-averaged information on state
mixtures in a set of four real parameters \[16\]. The four Stokes parameters are:

\[
S_0 = I = \langle S_{X/R} \rangle + \langle S_{Y/L} \rangle = |E_x|^2 + |E_y|^2 \\
S_1 = Q = \langle S_X \rangle - \langle S_Y \rangle = |E_x|^2 - |E_y|^2 \\
S_2 = U = \langle S_I \rangle - \langle S_L \rangle = \left[ |E_x E_y^*| + |E_x E_y| \right] \\
S_3 = V = \langle S_L \rangle - \langle S_R \rangle = -i \left[ |E_x E_y^*| - |E_x E_y| \right]
\] (4.12)

This array of equations is to be interpreted as follows: the first and second columns are two widely used notations for the Stokes parameters. The third column indicates their physical meaning: \( S_0 \) is the total time averaged signal \( \langle S(t) \rangle \), as it would be measured without any polarization filter. \( S_1 \) and \( S_2 \), being measured with linear filters, characterize the degree of linear polarization. \( S_3 \) becomes the degree of circular polarization. The fourth column of the array of equations expresses the parameters in the linearly polarized \( x,y \)-basis. All quantities already appeared in Eqs. 4.6-4.9, the only new aspect is the time averaging, denoted with the brackets \( \langle \ldots \rangle \). Only four measurements are needed to determine all four Stokes parameters. Two of the six measurements indicated in the third column of Eqs. 4.12 are superfluous. On the other hand, increasing the number of measurements will improve the accuracy.

To appreciate fully the power of the Stokes approach we consider the following derived quantities:

\[
m_T = \sqrt{s_1^2 + s_2^2 + s_3^2} \quad m_L = \sqrt{s_1^2 + s_2^2} \quad m_C = s_3
\] (4.16-4.18)

Here \( s_i \) are the normalized versions of the Stokes parameters, i.e. \( s_i = S_i/S_0 \). The quantity \( m_T \) can be interpreted as the total degree of polarization, or, the degree of polarization coherence. It is easy to verify that \( m_T = 1 \) for a coherent pure state with constant \( E_x \) and \( E_y \). In other words, for such pure states \( s_1^2 + s_2^2 + s_3^2 = 1 \). The quantities \( m_L \) and \( m_C \) characterize the degree of linear and circular polarization, respectively. Note that \( m_L^2 + m_C^2 = m_T^2 \). As an opposite example consider an incoherent mixture of linearly polarized states \( \langle e \rangle = (\cos(\psi) \sin(\psi)) \), where \( \psi \) is uniformly distributed between 0 and \( 2\pi \). In this case \( \langle E_x E_y^* \rangle = \langle \cos(\psi) \sin(\psi) \rangle = 0 \) and \( \langle E_x^2 \rangle = \langle E_y^2 \rangle \). Thus, \( m_T = 0 \). In general, \( m_T < 1 \) for partially polarized light.

### 4.3 General experiment design, Coherence matrix

If Stokes had lived 100 years later, he would certainly have defined his parameters in terms of Pauli matrices. It is quite easy to verify that:

\[
S_0 := \langle \langle E | \sigma_0 | E \rangle \rangle \quad S_1 := \langle \langle E | \sigma_1 | E \rangle \rangle \quad S_2 := \langle \langle E | \sigma_2 | E \rangle \rangle \quad S_3 := \langle \langle E | \sigma_3 | E \rangle \rangle
\] (4.19)

where

\[
\sigma_0 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad \sigma_1 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \quad \sigma_2 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad \sigma_3 = \begin{pmatrix} 0 & i \\ -i & 0 \end{pmatrix}
\] (4.20)

Warning!!!: Our notation of Pauli matrices differs slightly from the standard used in QT. Since optics and QT developed along separate paths, we had three options: i) a confusion of signs and subscripts. ii) different notation for Stokes parameters,
iii) different notation for the $\sigma$-s. Doing optics, we decided on iii). Pauli matrices exhibit certain peculiar and useful properties that can be easily understood when one realizes their relation with the projection operators from Eqs. 3.16, 3.17 and 3.18:

$$
\sigma_0 \equiv F_{X/R} + F_{Y/L} \quad \sigma_1 \equiv F_X - F_Y \quad \sigma_2 \equiv F_f - F_\perp \quad \sigma_3 \equiv F_R - F_L \quad (4.21)
$$

To demonstrate the practical usefulness of Pauli matrices, we return to Eq. 4.3, which relates the signal transmitted through a filter $|f\rangle$ with an input state $|E\rangle$. Introducing the time averaging, the measurable becomes

$$
\langle S_f \rangle = \langle \langle E | f \rangle \langle f | E \rangle \rangle = \langle \langle f | E \rangle \langle E | f \rangle \rangle \quad (4.22)
$$

Note here the projection operators $|f\rangle \langle f|$ or $|E\rangle \langle E|$; recall Eq. 3.15. Upon evaluating $|E\rangle \langle E|$ we find that Eq. 4.22 can be expressed in terms of a matrix $C$ as $\langle S_f \rangle = \langle f | C | f \rangle$, where

$$
C = \langle |E\rangle \langle E| \rangle = \left( \begin{array}{cc}
E_x & E_x^* \\
E_y & E_y^*
\end{array} \right) = \left( \begin{array}{cc}
\langle E_x E_x^* \rangle & \langle E_x E_y^* \rangle \\
\langle E_y E_x^* \rangle & \langle E_y E_y^* \rangle
\end{array} \right) \quad (4.23)
$$

Note that the complex hermitian “coherency matrix” $C$ contains the same information as the Stokes parameters $S$: on the diagonal there are two real numbers and the off-diagonal terms differ only in the sign of the imaginary part. Thus, four measurements are needed to determine $C$. Now we exploit a marvelous property of Pauli matrices: they form a complete set of independent matrices, i.e., any 2x2 matrix can be written as a linear superposition of $\sigma$-s. The reader may verify that the Pauli expansion of the coherency matrix $C$ reads:

$$
C = \frac{1}{2} \sum_{i=0}^{3} S_i \sigma_i \quad (4.24)
$$

Here $S_i$ are the Stokes parameters of the state $|E\rangle$! We insert this expansion into the right hand side of Eq. 4.22, observing thereby the quantities $\langle f | \sigma_i | f \rangle$. Thus, we recall Eq. 4.19 and obtain finally a beautifully simple expression for the measurable $\langle S_f \rangle$:

$$
\langle S_f \rangle = \frac{1}{2} \sum_{i=0}^{3} s_{fi} S_i \quad (4.25)
$$

Here $s_{fi}$ are the Stokes parameters of the filter state $|f\rangle$. Isn’t it nice?! The practical use of the small exercise with Pauli matrices for the polarization analysis should be obvious: Eq. 4.25 is only one linear equation for the four unknowns $S_0, S_1, S_2$ and $S_3$. Using four linearly independent measurements $\langle S_f \rangle$, one can determine all four Stokes parameters. If one is willing to do more measurements, one obtains an over-determined equation system, which can be solved by standard techniques, in order to improve the accuracy of the determination of the parameters. The required filter coefficients $s_{fi}$ are easily calculated upon programming the definitions 4.19 in Maple or Mathematica.

### 4.4 Stokes vectors

To conclude the story about Stokes parameters, we note that the sum in Eq. 4.25 can be interpreted as the scalar product of two 4D Stokes-vectors:

$$
\langle S_f \rangle = \frac{1}{2} \vec{s}_f \cdot \vec{S}, \quad \text{where} \quad \vec{S} = (S_0, S_1, S_2, S_3) \quad (4.26)
$$

30
The Stokes vectors of the three fundamental pairs of states can be easily obtained by inserting Eqs. 3.11 into 4.19:

\[
\begin{align*}
\vec{s}_X &= (1, +1, 0, 0) & \vec{s}_Y &= (1, -1, 0, 0) \\
\vec{s}_\| &= (1, 0, +1, 0) & \vec{s}\!\!\downarrow &= (1, 0, -1, 0) \\
\vec{s}_R &= (1, 0, 0, +1) & \vec{s}_L &= (1, 0, 0, -1).
\end{align*}
\] (4.27)

One may now realize why those three states are so fundamental: using these three pairs we can define a set of base vectors of the 4D Stokes space:

\[
\begin{align*}
\vec{s}_0 &= \frac{1}{2} (\vec{s}_X + \vec{s}_Y) = (1, 0, 0, 0) & \text{(4.28)} \\
\vec{s}_1 &= \frac{1}{2} (\vec{s}_X - \vec{s}_Y) = (0, 1, 0, 0) & \text{(4.29)} \\
\vec{s}_2 &= \frac{1}{2} (\vec{s}_\| - \vec{s}\!\!\downarrow) = (0, 0, 1, 0) & \text{(4.30)} \\
\vec{s}_3 &= \frac{1}{2} (\vec{s}_R - \vec{s}_L) = (0, 0, 0, 1) & \text{(4.31)}
\end{align*}
\]

The vector \(\vec{s}_0\) represents an unpolarized light field, such as would result from an incoherent superposition of any two orthogonal pure states with equal intensities. The remaining three can be called polarization unit vectors of the three fundamental states. Note, however, that the states \(\vec{s}_1, \vec{s}_2\) and \(\vec{s}_3\) do not carry any power. They must be always combined with \(\vec{s}_0\).
Chapter 5

Beam filters and beam states

The optical system we have in mind is not an array of polarizing elements and deflectors on an optical table but a sample of biological tissue. Consider the situation sketched in Fig. 5.1. The sample is irradiated by a nice coherent beam wave, but what comes out, apart from some specular reflection, is a disordered mixture of quasi-spherical waves. From this mixture one must first filter out a beam wave suitable for polarization analysis. One way to do such wave filtering is to place a very small detector very far in the far field, so that the superposition of the partial waves from the tissue appears as if coming from a point source, far enough to have a flat wave front over the entire detector area. In M. von Laue’s words: ".. we move the detector point always so far away, that if a spherical wave would originate in this point, it would look like a plane wave in the region where the scattering bodies are contained."[17]. One can show that the plane wave approximation requires that $L^2 \ll \lambda R$, where $R$ is the distance to the detector point and $L$ is the size of the region of interest. This far field assumption is certainly well fulfilled when observing stellar atmospheres, but hard to fulfill on a distance separating two consecutive scattering events during propagation in tissue. Anyway, this kind of wave filtering is highly impractical (and useless for imaging!), and therefore it is only used in textbooks. In optical practice one uses spatial filters consisting of lenses and apertures that select simultaneously the $k$-vector range and the observed position. This kind of wave filtering works fine, especially for imaging. But it is still not as perfect as polarization filtering with a Glan-Thompson prism: to select a pure beam state analogous to a pure polarization state one would have to use a nearly point-like detector giving nearly zero power. Glan-Thompson prism does the filtering virtually without any power loss.

While perfect polarizers have been around for more than a hundred years, single-mode beam filters are a relatively new addition to photonic instrumentation. A beam filter is constructed from a single-mode optical fiber (single-mode wave guide), a collimating lens and a polarization filter, as illustrated in Fig. 5.1. The beam filter works both ways, as a transmitter or a receiver. No matter what light source (e.g. a laser diode of appropriate wavelength) is coupled into the single-mode fiber, there is always the same nice beam emanating through the collimator lens. Replacing the laser diode with a detector, one obtains a single-mode receiver: no matter what light field is observed, one is sure to select always the same nice beam that propagates into the fiber. Besides being a nice practical device for scattering experiments [18, 19], the single-mode receiver has also a theoretical advantage: using the single-mode concepts one can treat scattering problems (involving interference) in a rigorous yet rather simple fashion [20]. Beam filters, as theoretical concepts and/or practical devices, will be useful for polarization imaging (Sec. 11.6), for the detection of scattered light (Sec. 7.1) as well as for
Figure 5.1: Random light field generated by scattering in a tissue sample and pure beam selection by a single-mode receiver. Note that we avoid specular reflection.

bridging electrodynamics with Monte Carlo simulations.

5.1 Mathematical representation of beam filters

Electrodynamic theory provides tools to make the mode selection by a single-mode receiver very explicit: One can define an electromagnetic field \( \hat{\mathbf{E}}^o, \hat{\mathbf{B}}^o \equiv |o\rangle \) that possesses all the properties of the classical electromagnetic field which would emanate from a transmitter (recall von Laue’s remark!), except that its propagation is reversed (time reversal) and that the field is empty. We shall call this field observation mode. The observation mode \( \hat{\mathbf{E}}^o, \hat{\mathbf{B}}^o \) is a virtual field in the original sense of the word: a virtual field does not carry photons, but it has the potential to do so. (In this spirit, one may regard all fields as purely geometrical constructions that provide a path system for the transport of particles.) In chapters 3 and 4 we pointed out that filtering is synonymous with the projection operation. Projection, in turn, requires the existence of a norm \( \langle f | f \rangle \) and of a scalar product \( \langle E | f \rangle \). Recall that the norm of a beam state is defined as

\[
\langle E | E \rangle = \frac{1}{4} \frac{1}{\mu_0} \int_S [\mathbf{E} \times \mathbf{B}^* + \mathbf{E}^* \times \mathbf{B}] \cdot \hat{\mathbf{n}} \, d^2r. \tag{5.1}
\]

where \( A \) is an arbitrary surface enclosing the source of the beam (i.e. a laser); \( \hat{\mathbf{n}} \) is the outer normal to this surface. In the case of an ordinary laser beam, as defined in Eq. 2.29, or in the quasi plane beam wave version in Eq. 2.38, the norm is the power carried in the beam: \( \langle E | E \rangle = P \). A virtual observation beam does not carry any power, but otherwise it looks much like an ordinary beam wave. We are mostly interested in polarized paraxial beams in the Rayleigh range, whose electric field is given by Eq. 2.38. Since the virtual observation mode does not carry any power, we define the observation beam as a “powerless” version of Eq. 2.38:

\[
\hat{\mathbf{E}}^o(\mathbf{r}) = \ldots \sqrt{\frac{2}{c\varepsilon_0}} X(\mathbf{r}) \hat{\mathbf{e}}_o e^{i k \hat{\mathbf{s}}_o \cdot \mathbf{r}} \tag{5.2}
\]

The profile \( X(\mathbf{r}) \) is the beam profile from Eq. 2.36, expressed, however, in an arbitrary oriented coordinate system. A corresponding transformation concerns also the components of the polarization vector \( \hat{\mathbf{e}}_o \) and the propagation vector \( \hat{\mathbf{s}}_o \) of the observation beam. Upon inserting this powerless \( \hat{\mathbf{E}}^o \) into Eq. 5.1, one finds that

\[
\langle o | o \rangle = \frac{1}{4} \frac{1}{\mu_0} \int_S [\hat{\mathbf{E}} \times \hat{\mathbf{B}}^* + \hat{\mathbf{E}}^* \times \hat{\mathbf{B}}] \cdot \hat{\mathbf{n}} \, d^2r = 1. \tag{5.3}
\]

33
Note that the “hat” mark over the field amplitudes does not mean that $\hat{E}$ and $\hat{B}$ were unit vectors. It is just a reminder of the fact that the norm of a virtual state $|o\rangle$ is 1, just as the norm of a unit Euclidean vector. Consequently, we define the scalar product of an external field $|E\rangle$ with an observation mode $|o\rangle$ as

$$\langle o|E \rangle = \frac{1}{4 \mu_0} \int_A [E \times \hat{E}^o + \hat{E}^o \times \hat{B}] \cdot \hat{n} \, d^2 r.$$  \hspace{1cm} (5.4)

where $A$ is an arbitrary surface enclosing the sink of the observation beam (a detector) but excluding the source of $|E\rangle$; $\hat{n}$ is the inner normal to this surface. If both fields are transversal (in the far field or in Rayleigh range) and their propagation vectors $\hat{s}_e$ and $\hat{s}_o$ nearly collinear, then the following expression is a useful approximation of 5.4:

$$\langle o|E \rangle \approx \frac{c \varepsilon_0}{4} \int_A \left( \hat{E}^o \cdot E \right) [\hat{s}_e + \hat{s}_o] \cdot \hat{n} \, d^2 r.$$  \hspace{1cm} (5.5)

Note here the Fresnel inclination factor $[\hat{s}_e + \hat{s}_o] \cdot \hat{n}/2$ and the polarization factor $\left( \hat{E}^o \cdot E \right)$, both well known from the rigorous diffraction theory. Eq. 5.5 is particularly useful for the description of the coupling of laser beams into single-mode fibers. We shall need it in the context of polarization imaging.

In some cases, the external field $|E\rangle$ is too complicated for an explicit expression, but the source of this field, i.e. the oscillating current distribution $j$ that radiates this field, is simple. For this case the following (exact) version of 5.4 can be derived using a magic wand [21] from the tools of electromagnetism (Poynting theorem from Eq. 2.8 and reciprocity theorem [22]):

$$\langle o|E \rangle = -\frac{1}{4} \int_V \hat{E}^o \cdot j \, d^3 r.$$  \hspace{1cm} (5.6)

Here $V$ is an arbitrary volume that encompasses the source $j$ of $|E\rangle$ but excludes the sink of $|o\rangle$. Equation 5.6 will be quite useful for concise treatment of scattering problems that involve interference, in particular in the single scattering limit.

Knowing the scalar product, we can calculate the field in the beam that is selected by the receiver as the projection of the external field on the receiver mode

$$|E^o\rangle = |o\rangle \langle o|E \rangle = |o\rangle \mathcal{E}$$  \hspace{1cm} (5.7)

The abbreviation $\mathcal{E} \equiv \langle E|o \rangle$ suggests to interpret the scalar product as a *complex amplitude*. In order to get the real field, one simply takes the virtual field from Eq. 5.2 and multiplies it with the complex amplitude $\mathcal{E}$:

$$E^o(r) = \hat{E}^o(r) \mathcal{E}.$$  \hspace{1cm} (5.8)

Note that $|o\rangle$ and $\langle o|\rangle$ are complex conjugate and thus the phase of the receiver mode does not enter the selected field $E^o(r)$. The selected field $E^o(r)$ inherits the phase of the source. Thus we can write the complex amplitude as

$$\mathcal{E} \equiv \langle E|o \rangle = \sqrt{P_{os}} e^{i\Phi_s},$$  \hspace{1cm} (5.9)

where $\Phi_s$ is the inherited phase. The received signal $P_{os}$, i.e. the power measured with a single-mode receiver, is given by the general expression, Eq. 4.3. Since the observation beam is normalized, we can write

$$P_{os} = |\langle E|o \rangle|^2 = |\mathcal{E}|^2$$  \hspace{1cm} (5.10)
Note that $P = |\mathcal{E}|^2$ is completely analogous with $I = |E|^2$. In many situations one does not care about the unmeasurable local details of the electromagnetic field of a beam wave and so the knowledge of $\mathcal{E}$ is all one needs. However, once $\mathcal{E}$ is known, we can calculate the field of the selected beam at any position, in particular in the far field, according to Eq. 5.8. For the calculation of $\langle E|\hat{o}\rangle$ one can use any of the expressions from Eqs. 5.4, 5.5 or 5.6. Of course, one would choose the form which is most convenient for a particular problem.

5.2 Virtual beams as base vectors

In chapter 3.3 we introduced the Jones formalism in a rather careless traditional way. We did not care about the details of the structure of the fields, essentially we pretended working with plane waves, where the field amplitude $E$ and the polarization vector $\hat{e}$ is all what matters. In the laboratory practice, however, one usually works with beams. The measured quantity is the beam power $P$ and not some local intensity $I(\mathbf{r}) \sim |E(\mathbf{r})|^2$. Having understood the concept of virtual modes as unit vectors in the space of the solutions of Maxwell Equations, we are now in the position to mend the inconsistency and put the Jones formalism on sound formal footing. To express a laser beam in, for example, XY-polarized base, we write

$$\mathbf{E}^l(\mathbf{r}) \equiv |E^l\rangle = |X\rangle \mathcal{E}_x + |Y\rangle \mathcal{E}_y, \quad (5.11)$$

where $|X\rangle$ and $|Y\rangle$ are unit vectors, i.e. virtual beams normalized such that $\langle X|X \rangle = 1$ and $\langle Y|Y \rangle = 1$. Explicit definitions of the two unit vectors are

$$|X\rangle \equiv \hat{E}^x(\mathbf{r}) = \cdots \sqrt{\frac{2}{c\epsilon_0}}X(\mathbf{r}) \hat{x} e^{i k \hat{s}_l \cdot \mathbf{r}} \quad (5.12)$$

$$|Y\rangle \equiv \hat{E}^y(\mathbf{r}) = \cdots \sqrt{\frac{2}{c\epsilon_0}}X(\mathbf{r}) \hat{y} e^{i k \hat{s}_l \cdot \mathbf{r}}$$

As an exercise one may use Eq. 5.5 to verify that the complex amplitudes $\mathcal{E}_X$ and $\mathcal{E}_Y$ are the scalar products defined in previous chapter. A Jones vector is defined as a two dimensional vector whose components are the complex amplitudes

$$|E\rangle := \begin{pmatrix} \mathcal{E}_x \\ \mathcal{E}_y \end{pmatrix} \quad \langle E\rangle := \begin{pmatrix} \mathcal{E}_x^* \\ \mathcal{E}_y^* \end{pmatrix}, \quad (5.13)$$

so that

$$\langle E|E \rangle = |\mathcal{E}_x|^2 + |\mathcal{E}_y|^2 = P \quad (5.14)$$

The virtual beam waves propagate through the optical system just as ordinary waves, which would allow us to extend Jones formalism to optical systems of the “more familiar type involving lenses, prisms, etc.” (Recall the citation from the preamble of Sec. 3.)
Chapter 6

Polarization analysis of an unknown optical system

6.1 Designing the experiment, from Jones to Perrin-Mueller matrix

Having filtered out a beam wave from the scattered, transmitted or reflected light we can concentrate on polarization analysis. The geometry is sketched in Fig. 6.1. Note that we are trying to avoid specular reflection, because we would like to obtain information from the depth of the sample. The illumination and the observation channels are equipped with polarization optics as needed to manipulate and filter the polarization states. The analysis is done in the \( \pi \)-base system, recall Sec. 3.5. How much information is contained in the polarization of the scattered light field and what measurements are needed to retrieve this information? After the exercise in chapters 3.3 and 4, the answers to these questions should not be too difficult. The transformation of the input state \( |e^{in}\rangle \) into an output state \( |E^{out}\rangle \) is usually achieved by an amplitude transfer matrix (Jones matrix) \( T \), which, however, depends on the geometry of illumination and observation:

\[
|E^{out}(t)\rangle = T(\hat{s}_i, \hat{s}_o, t) |e^{in}\rangle ,
\]  

(6.1)

Note that in Eq. 6.1 we are using two different symbols for the input and for the output: whereas \( |e^{in}\rangle \) is a coherent state normalized so that \( \langle e^{in}|e^{in}\rangle = 1 \), the output
state $|E_{out}(t)\rangle$ is in general a mixture. In fact, $T(\hat{s}_i, \hat{s}_o, t)$ exhibits random fluctuations, unless the tissue sample is stiffly frozen. But even if the sample were frozen, we are not as much interested in a particular random realization of the tissue structure as in an ensemble average characteristic for the given tissue type. The characterization will always involve averaging of the filtered signals.

We return to Eq. 4.22, inserting therein the output state $|E_{out}(t)\rangle$ from Eq. 6.1:

$$\langle S_{fe}(T) \rangle = \langle \langle f | E \rangle \langle E | f \rangle \rangle = \langle \langle f | T | e \rangle \langle e | T^\dagger | f \rangle \rangle$$  \hspace{1cm} (6.2)

On the right hand side $\dagger$ means Hermitian adjoint, i.e. transposed conjugate of $T$. Obviously, any measured $\langle S_{fe}(T) \rangle$ is some real linear combination of the correlations $\langle T_{ij} T_{k\ell}^\dagger \rangle$. Thus, the 16 averages $\langle T_{ij} T_{k\ell}^\dagger \rangle$ represent the full amount of information on the observed system which can be retrieved from time averaged observations. There are 4 real quantities $\langle T_{ij} T_{ij}^\dagger \rangle$, the remaining 12 are complex conjugate pairs. Thus, the information can be retrieved using 16 linearly independent combinations of the filters $|e\rangle$ and $|f\rangle$. The problem is completely analogous with the characterization of an unknown state vector $|E\rangle$, but organizing the 16 measurements in a clever way is a demanding task! Luckily we prepared the tools: In chapter 4.3 we learned how to re-write $\langle \langle f | E \rangle \langle E | f \rangle \rangle$ in terms of a scalar product of two Stokes vectors: $\langle S_{fe} \rangle \sim \vec{s}_f \cdot \vec{S}$, where $\vec{S}$ is the Stokes vector of the output state $|E\rangle$. If we were able to find a 4x4 matrix $\mathbf{M}$ which transforms the Stokes vector $\vec{s}_e$ of the input state $|e\rangle$ into the output vector $\vec{S}$, then we could write $\langle S_{fe} \rangle \sim \vec{s}_f \cdot \mathbf{M} \cdot \vec{s}_e$, which would be quite nice. Such a matrix $\mathbf{M}$ does indeed exist and is called Mueller matrix, despite the fact that it was introduced by the French physicist Francis Perrin (son of Jean Perrin) already in 1942 [23]. To derive the Perrin-Mueller matrix, we shall exploit the wonderful properties of Pauli matrices. First we extract from Eq. 6.2 the identity $|E\rangle \langle E| \equiv T |e\rangle \langle e| T^\dagger$. (For the sake of readability we omit for a while the outer brackets of time averaging.) Expanding the projectors on both sides in terms of Pauli matrices we obtain a matrix equation that relates the Stokes parameters of the input and output:

$$\sum_k S_k \sigma_k = \sum_i s_{ei} T \sigma_i T^\dagger$$ \hspace{1cm} (6.3)

These equations look a little bit strange, but they can be converted into a set of four ordinary linear equations using the following properties of the Pauli matrices: i) the matrices $\sigma_1$, $\sigma_2$ and $\sigma_3$ are traceless, i.e. the sums of their diagonal elements are always zero. Only the identity matrix $\sigma_0$ has a non-zero trace, namely 2. ii) the product of two Pauli matrices $\sigma_k \sigma_j$ is always a traceless Pauli matrix, except for $\sigma_j \sigma_j = \sigma_0$. In other words $\text{Trace}(\sigma_k \sigma_j) = 2\delta_{k,j}$, where $\delta_{k,j} = 1$ if $j = k$ and $\delta_{k,j} = 0$ if $j \neq k$. Thus, upon multiplying both sides of Eq. 6.3 with $\sigma_j$ and carrying out the Trace operation we obtain:

$$S_j = \sum_i s_{ei} M_{ji}, \quad \text{where} \quad M_{ji} = \frac{1}{2} \text{Trace}(T \sigma_i T^\dagger \sigma_j)$$ \hspace{1cm} (6.4)

Here $M_{ji}$ are the elements of the Mueller matrix, $j$ is row. Since the matrix elements can be easily calculated using an algebra program, we refrain from reproducing the explicit results. The expressions look somewhat ugly (see e.g. [24]), but the formal beauty of the Perrin-Mueller approach becomes apparent in the following expression for the measured signal

$$\langle S_{fe} \rangle = \frac{1}{2} \vec{s}_f \cdot \mathbf{M} \cdot \vec{s}_e.$$ \hspace{1cm} (6.5)

Upon carrying out the multiplications in Eq. 6.5 for one input state $\vec{s}_e$ and one filter $\vec{s}_f$, one obtains one of the 16 equations, which are needed to determine the 16 elements of
an unknown Mueller matrix from the measurements \( \langle S_{f_e} \rangle \). One may rightly guess that the sixteen measurements can be done by permuting the special states \(|X\rangle, |Y\rangle, |/\rangle\) and \(|R\rangle\), one of those elements in the illumination channel and one in the observation. Of course these 4x4 measurements are not the only way how to setup the equation matrix. A little bit of patience will improve the accuracy as well as the understanding.

We suggest to do at least 6x4 measurements: the three fundamental pairs for the input, and at least 4 measurements to determine each of the output Stokes vectors \( \vec{O}_e = \vec{M} \vec{s}_e \). The Stokes vectors \( \vec{s}_e \) of the three fundamental pairs are given in Eq. 4.27. One immediately sees that

\[
\begin{align*}
\vec{O}_X &= \vec{M}_0 + \vec{M}_1 \\
\vec{O}_Y &= \vec{M}_0 - \vec{M}_1 \\
\vec{O}_R &= \vec{M}_0 + \vec{M}_3 \\
\vec{O}_L &= \vec{M}_0 - \vec{M}_3,
\end{align*}
\]

where \( \vec{M}_j \) are the columns of the Mueller matrix \( \vec{M} = (\vec{M}_0, \vec{M}_1, \vec{M}_2, \vec{M}_3) \). Thus, by adding and subtracting the Stokes outputs of the fundamental pairs, one determines entire columns of the Mueller matrix. A nice compilation of other different possibilities to organize the measurements, as well as practical hints, can be found in [25]. In fact, there are infinitely many possibilities. Contemporary computation tools allow to solve an arbitrary large set of over-determined equations and good algorithms complain reliably if the system is under-determined (SVD). The bottleneck is the measurement. As with the determination of an unknown Stokes vector, the more measurements one is willing to do, the more accurate is the result.

### 6.2 Interpreting the Perrin-Mueller matrix, alternative codings

Having understood the column structure of the Perrin-Mueller matrix, we can attempt an interpretation. This can be best done by applying the matrix to linear combinations of the polarization unit vectors defined in Eq. 4.28. The unpolarized unit vector \( \vec{s}_0 \) picks the column \( \vec{M}_0 \). This column represents re-polarization of an unpolarized state mixture: if any of the elements \( M_{k0}, k = 1, 2, 3 \) is non zero then the output is at least partially polarized. (Recall for example the polarization of sunlight by scattering in the atmosphere.) The remaining three columns (together with column 0) describe the transfer of the power from one of the polarization states to the others. Note, for example, that in order to make circular polarization out of linear states \( \vec{s}_X \) or \( \vec{s}_s \), the elements \( M_{30}, M_{31} \) and/or \( M_{32} \) must be non-zero. For the elements of the amplitude transfer matrix \( \vec{T} \) this means either non-zero off-diagonal terms \( T_{xy}, T_{yx} \) so that e.g. \( \Im \langle T_{xx} T_{yx}^* \rangle \neq 0 \) (in \( M_{30} \) and \( M_{31} \)), or complex diagonal terms with phase shift between \( T_{xx} \) and \( T_{yy} \) so that \( \Im \langle T_{xx} T_{yy}^* \rangle \) (in \( M_{32} \)). On the other hand, to make linear polarization out of \( \vec{s}_R \) or \( \vec{s}_L \) it is sufficient that \( \langle |T_{xx}|^2 \rangle \neq \langle |T_{yy}|^2 \rangle \) (in \( M_{10} \)).

Apparently, the information contained in \( \vec{M} \) is not as easily interpretable as in the case of the Stokes vector. It may well be that this particular coding rather obscures the sought diagnostic information instead of enhancing it. Therefore, instead of brooding too long over \( \vec{M} \), one should always try an alternative coding. For example, one can survey the real and imaginary parts of the elements \( \langle T_{xy} T_{yx}^* \rangle \) of the 4D coherence matrix, can be arranged in a real matrix \( \vec{P} \), and which contain the same information as the Mueller matrix. Finally, one can survey directly the matrix \( \vec{S} \) containing the 16 independent observables \( \langle S_{f_e} \rangle \) instead of converting \( \vec{S} \) into the \( \vec{P} \) or \( \vec{M} \) matrices. For example, without birefringence it is difficult to make circular polarization out of
linear. Thus the measurements $\langle S_{RX} \rangle$ and/or $\langle S_{LX} \rangle$ may be suited to diagnose birefringence. We have at least three matrices, $S$, $P$, and $M$, which contain exactly the same information, only in a different way. Which coding would provide the most discernable signature of, say, the chirality of the medium? It would be nice to find a set of 16 really existing or imaginable optical devices (filters, blockers, depolarizers, transformers, attenuators, ...) whose 16 matrices would form a complete set for an intelligible decomposition of $S$, $P$ or $M$ matrices. In any case, when designing data analysis procedures for tissue diagnostics one should be guided not only by canonical formalisms but also by scientific imagination. The research on the interpretation of the Mueller and other matrices keeps continuing [26].

### 6.3 Mueller versus Jones

Upon carrying the reasoning from the previous chapter a little bit further, one discovers that by adding new elements into the optical path the game with Mueller matrices and Stokes vectors can be continued. Propagation of polarized light through the optical system can be analyzed with the so-called Mueller formalism, i.e. in terms of Mueller matrices and Stokes vectors, as easily as with the Jones formalism which we employed in Sec. 3. Any Jones matrix can be converted into its Mueller matrix using the trace-formula in Eq. 6.4, and any Jones vector $|E\rangle$ can be converted into a Stokes vector using Eqs. 4.19. Thus the following two expressions are almost completely equivalent:

$$\vec{S}_{\text{out}} = M_F M_{\lambda/2} R \vec{S}_{\text{in}}$$

Thus, we have at least two possibilities how to rationalize the work on propagation of polarized light through biological tissues: i) Jones calculus, that operates with 2x2 complex matrices and propagates the amplitudes, and, ii) Stokes-Mueller calculus, that operates with 4x4 real matrices and propagates the probabilities. Hans Müller, an MIT scientist, published his calculus in 1948, only a few years after the harvardian Jones ([27]). As it tends to happen between MIT and Harvard, the two approaches are often felt as a sort of competition, which manifests itself in the following statement that is found obligately in most texts on the polarization: “Jones formalism can’t treat unpolarized light”. In our opinion, this is a misunderstanding. There is no such thing as “unpolarized light”. But there are light fields where the polarization vector is a randomly fluctuating quantity. One can always do Jones calculus and calculate probabilities or averages at any step of the light path. In a way the Mueller-matrix characterizes an ensemble of Jones matrices, as has been proven formally by Kim, Mandel and Wolf [28], who felt a need to “reconcile” Harvard and MIT.

We share here the opinion of those pragmatic scientists, who regard the two algorithms as useful complements. Jones’ approach is certainly more fundamental, i.e. closer to the underlying quantum world. However, it is a little bit tedious in usage when it comes to the ensemble averaged observables. This is because the observables are quadratic in field amplitudes, but it is a pain to keep together the pairs of amplitudes to be averaged as a product. Here Stokes, Perrin and Müller come in: in the Stokes vectors and Perrin-Mueller matrices such products are neatly organized, ready to be averaged. But there is a price to be paid: the common phase of the two components of the state vector is gone. Thus, Stokes-Mueller calculus is suited for problems, where amplitude interference between different photon paths can be neglected, or is willingly neglected. Chandrasekhar’s treatment of polarized radiative transfer is the
most prominent example [14]: the linear relationships between the probabilities greatly simplify the complicated analytics. On the other hand, the most interesting and relevant problems require numerical calculations, and here the difference between Jones and Müller has faded: today’s computing hardware and software handles 2x2 complex matrices as easily as 4x4 real matrices. In fact, it takes 64 multiplications to calculate a product of two Perrin-Mueller matrices as compared with only 32 multiplications in a product of two Jones matrices.

We shall employ the Jones approach for the development of polarized scattering and polarized Monte Carlo simulations. This amplitude based formalism is closer to the fundamental physics, preserves the phase and allows thus to include path interference. Such interference effects are exploited by several practical analytical techniques, such as imaging speckle interferometry [29] or diffusing wave spectroscopy [30]. These techniques are beyond the scope of the present book, but we would like to keep the path open for future developments. However, we shall use the Stokes-Perrin-Müller concepts (or S, P-matrix concepts) when ever appropriate, that is when it comes to averaging and organizing the measured observables in a rational and efficient way.
Chapter 7

Basic scattering concepts, Born approximation

Biological tissues are highly heterogeneous condensed media, whose optical properties are quite difficult to model. In fact, it is much easier to say what tissues are not: they are not dilute suspensions of independently scattering sphere-like particles. This statement sounds much like a truism, which it is. However, the mathematics employed for the description of light propagation in tissues were originally developed for propagation of incoherent stellar light in interstellar atmospheres [14], which certainly are such particle suspensions, even extremely dilute. Surprisingly, these techniques are highly successful in propagating coherent laser light in tissues. It seems that the suggestive power of this success strongly influenced the way how we think about tissue optics: when thinking about scattering in tissues, one inevitably has Mie particles in mind. It was only in the last decade that new approaches to tissue optics began to emerge. Instead of picturing the tissue as a cloud of independent particles, one seeks to characterize its random dense structure in terms of density correlation functions or spatial power spectra [31]. These developments appear to have drawn their inspiration from two sources. One is the wave propagation in turbulent atmosphere, as discussed e.g. in [32], and the second inspiration are concepts from statistical physics that were originally developed for small angle X-ray and neutron scattering [31, 33, 34, 35] in soft condensed matter. Since soft condensed matter seems to be much closer to biological tissue than turbulent atmosphere, we shall pursue the second path. In small angle scattering one usually employs two assumptions: i) the interaction of the radiation with the matter is weak, so that the scattering can be treated in the first Born approximation; multiple scattering is usually negligible. ii) because of the inherently small scattering angle (typically 0.1 - 5°), polarization can be neglected. Fortunately, we can retain the first part of assumption i), because the refractive index differences of the tissue constituents are relatively small. They lay between the extremes water and fat, i.e. in the range 1.33-1.45 [36]. However, multiple scattering is what we are finally interested in. Moreover, the present task is to include in the multiple scattering the effect of polarization. Thus, in order to make use of the new approaches to tissue optics, we must learn to combine the concepts involving the correlation functions with the polarization concepts, including the Jones matrix of the scattering process.

7.1 Scattering geometry

We begin with the most simple and well known case of scattering from a single dielectric particle, for example a single molecule. The particle scatters laser light: The lasers
oscillating electric field $E_l(r) \exp(-i\omega t) \equiv |E_l\rangle$ excites in the particle an oscillating current density $j(r, t) = j(r) \exp(-i\omega t)$, which in turn radiates electromagnetic energy through a quasi-spherical wave $E^s(r) \exp(-i\omega t) \equiv |E^s\rangle$. (Dirac’s bra-ket-notation was introduced in Chap. 3.3.) From the scattered field that radiates in all directions we select for observation a certain component $|E^os\rangle$. The selection is done with a beam filter, as discussed in Chapter 5 and illustrated in Fig. 5.1. Scattering by isolated particles is well known from standard texts, but we use the opportunity to prepare tools for treating more difficult aspects of light scattering and propagation in tissues. What we are aiming at is a scattering transition matrix, i.e. a Jones matrix $T$ that captures the transition of a photon from the laser beam into the receiver beam, including not only polarization and amplitude of the process but also the phase. The matrix $T$ operates on the Jones vectors of complex amplitudes (recall Sec. 5.2):

$$
\begin{pmatrix}
E^L_x \\
E^L_y
\end{pmatrix}
= T_{lo}^L
\begin{pmatrix}
E^O_x \\
E^O_y
\end{pmatrix}
\quad (7.1)
$$

The superscripts refer to the fact that the Jones vectors are defined in two different coordinate systems, L and O. We need such transition matrix in order to design a model for polarized Monte Carlo simulation.

The geometry of the experiment is sketched in Fig. 7.1. The tripod $\mathbf{\hat{x}}_L, \mathbf{\hat{y}}_L,  \mathbf{\hat{z}}_L$ defines the laboratory coordinate system with the origin at $r_L$. The gray horizontal plane symbolizes desktop of an optical table as a plane spanned by the unit vectors $\mathbf{\hat{x}}_L, \mathbf{\hat{y}}_L$. The optical axis of the illuminating laser beam $\mathbf{\hat{s}}_i$ coincides with $\mathbf{\hat{z}}_L$. The $\mathbf{\hat{x}}_L$-axis is perpendicular to the optical table onto which the elements needed to manipulate the polarization state of the laser are mounted. The molecule is expected to scatter in all directions and therefore the receiver and the elements for polarization analysis are mounted on a rotational stage, which allows to set the observation angles $\phi$ and $\theta$. The laser beam and the receiver beam are aligned so that they intersect in a common focus at $r_f$. The propagation vectors $\mathbf{\hat{s}}_i$ and $\mathbf{\hat{s}}_o$ define the scattering plane $\pi$, as discussed in Sec. 3.5. The molecule is somehow fixed close to the focus $r_f$, in the region where the two beams overlap. Details of this region are shown in Fig. 7.2. Since positioning of single molecules is rather difficult, we allow for a small misalignment so that the position of the molecule is $r_m = r_f + \mathbf{m}$. Note that there are three coordinate systems involved. In addition to the lab system there are the two coordinate systems: the I-system $\mathbf{\hat{x}}, \mathbf{\hat{y}}, \mathbf{\hat{s}}_i$ with the origin at $r$, and the O-system $\mathbf{\hat{x}}, \mathbf{\hat{y}}, \mathbf{\hat{s}}_o$ with the origin at $r_o$. 

Figure 7.1: Scattering geometry. Details of the region around beam focus $r_f$ are shown in Fig. 7.2.
7.2 Transition matrix

Now we put the beam filter into operation in order to select from the scattered field \( |E^s⟩ \) an observed partial beam \( |E^os⟩ \) that propagates into the receiver. The selection is conveniently achieved using the projection technique discussed in chapter 5. In the absence of a polarizer in the detection path, the beam filter accepts two modes with the same beam profile but with orthogonal polarizations. Thus, the selected and observed component of the scattered field is the superposition of two orthogonal components, i.e. \( |E^os⟩ = |E^os_x⟩ + |E^os_y⟩ \), where

\[
\begin{align*}
|E^os_x⟩ &= |o_x⟩ \mathcal{E}_x^O = |o_x⟩ \langle o_x|E^s⟩ \\
|E^os_y⟩ &= |o_y⟩ \mathcal{E}_y^O = |o_y⟩ \langle o_y|E^s⟩ 
\end{align*}
\]

Once \( \mathcal{E}_x \) and \( \mathcal{E}_y \) are known, then the actual field \( \mathbf{E}^os(r) \) that is selected by the receiver from the scattered field can be obtained in any position \( r \) simply by combining the amplitudes with the virtual observation modes \( |o_x⟩ \) and \( |o_y⟩ \). Thus, the Jones vector consisting of the two amplitudes is a complete representation of the selected state:

\[
\begin{pmatrix}
\mathcal{E}_x^O \\
\mathcal{E}_y^O 
\end{pmatrix} = \begin{pmatrix}
\langle o_x|E^s⟩ \\
\langle o_y|E^s⟩ 
\end{pmatrix}
\]

Since the source of the scattered field, the current density distribution \( \mathbf{j}(r,t) \), is well localized within a small particle, we calculate the scalar products \( \langle o_x|E^s⟩ \) and \( \langle o_y|E^s⟩ \) using Eq. 5.6, reproduced here for convenience:

\[
\langle o|E^s⟩ = -\frac{1}{4} \int_V \mathbf{E}^os \cdot \mathbf{j} \, d^3r.
\]

The integration extends over the volume \( V \) of the particle. The only assumption that we make about the distribution \( \mathbf{j}(r,t) \) is that it results from linear dielectric response of the particle material to the impinging electric field (we neglect magnetic response). Recall from macroscopic electrodynamics that \( \mathbf{j}(r) = \nabla \mathbf{P}(r) / \partial t \) and \( \mathbf{P}(r) = \mathbf{\epsilon}_o \mathbf{X}(r) \mathbf{E}_{int}(r) \), where \( \mathbf{\chi}(r) \) is the susceptibility tensor and \( \mathbf{E}_{int}(r) \) is the local internal field, as it would be measured within the matter at the position \( r \). Here we employ a generalized version of this relation, which explicitly considers non-local response:

\[
\mathbf{j}(r,t) = -i\omega \int_V \mathbf{\epsilon}_o \mathbf{X}(r, r', \omega) \cdot \mathbf{E}'(r', t) \, d^3r'.
\]

Here \( \mathbf{\chi}(r, r', \omega) \) is a tensor density, a non-local version of the susceptibility tensor \( \mathbf{\chi} \). Note that \( \mathbf{E}'(r', t) \) is the impinging field as it would exist in the medium in the absence of the scattering particle, and not the internal field \( \mathbf{E}_{int}(r, t) \), that would be measured within the dielectric particle\(^1\). Note also that Eq. 7.5 is exact, no approximations are yet involved. The purpose of this purely formal expression is to make the vectorial linear relationship between the current density \( \mathbf{j} \) and the impinging field \( \mathbf{E}' \) explicit. Thus, the scalar product has the general form

\[
\langle o|E^s⟩ = \frac{i\omega}{4} \int_V \mathbf{E}^os(r) \cdot \mathbf{\epsilon}_o \mathbf{X}(r, r', \omega) \cdot \mathbf{E}'(r', t) \, d^3r' \, d^3r = \langle o|D|E'⟩. \tag{7.6}
\]

\(^1\)The internal field \( \mathbf{E}_{int}(r, t) \) is related to \( \mathbf{E}'(r', t) \) through the relation \( \mathbf{\chi}(r) \cdot \mathbf{E}_{int}(r, t) = \int_V \mathbf{\chi}(r, r', \omega) \cdot \mathbf{E}'(r', t) \, d^3r' \). This is essentially a short hand version of the Ewald Oseen extinction theorem, see e.g. [37]. Internal fields are the subject of classical treatments of scattering by particles, see e.g. Sec. 4.3. in [24].
On the right hand side we merely introduced a convenient abbreviation. Now we recall from chapter 5.2 that the impinging laser field, too, can be expressed as a linear superposition of two base states with two complex amplitudes $\mathcal{E}$. For compatibility reasons we chose XY- base states in the I-system and write:

$$|E^I\rangle \equiv \mathbf{E}^I(\mathbf{r}) = |l^I_x\rangle \mathcal{E}^I_{lx} + |l^I_y\rangle \mathcal{E}^I_{ly} \quad (7.7)$$

Applying this decomposition to the scalar products $\langle o|E^I\rangle$ (which are linear in $|E^I\rangle$), we may re-write Eq. 7.3 as

$$\left(\begin{array}{c} \mathcal{E}^O_x \\ \mathcal{E}^O_y \end{array}\right) = \left(\begin{array}{cc} \langle o_x|D|l^I_x\rangle & \langle o_x|D|l^I_y\rangle \\ \langle o_y|D|l^I_x\rangle & \langle o_y|D|l^I_y\rangle \end{array}\right) \left(\begin{array}{c} \mathcal{E}^I_{lx} \\ \mathcal{E}^I_{ly} \end{array}\right) \quad (7.8)$$

Right hand side of Eq. 7.8 defines the desired transition matrix in the I-system. The transition from the L-system to the I-system is achieved by a simple 2D rotation $\mathbf{R}_2(\phi)$:

$$\mathbf{T}^L_{lo} \equiv \mathbf{T}^I_{lo} \mathbf{R}_2(\phi) = \left(\begin{array}{cc} \langle o_x|D|l^I_x\rangle & \langle o_x|D|l^I_y\rangle \\ \langle o_y|D|l^I_x\rangle & \langle o_y|D|l^I_y\rangle \end{array}\right) \left(\begin{array}{cc} \cos \phi & \sin \phi \\ -\sin(\phi) & \cos(\phi) \end{array}\right) \quad (7.9)$$

Equation 7.9 is exact and quite general. It applies to any kind of linear scattering system, illuminated by any kind of laser field and observed with any kind of single mode receiver. The real work begins with the calculation of the matrix elements $\langle o|D|l\rangle$.

### 7.3 Matrix elements

The first step is to specify the impinging laser field and the observation modes. The laser beam is a quasi-monochromatic quasi-plane wave as it was introduced in Sec. 2.3.2. The laser power is $P_l$. The beam profile close to focus is $X_l(\mathbf{r})$, the propagation is directed along $\hat{s}_l$ and the polarization is specified by the complex polarization vector $\hat{e}_l$ (with components $e^L_{xl, yl, 0}$ in the L-system). We also specify the phase $\Phi(i)$ in a reference plane at $\mathbf{r}_l$ and $t = 0$:

$$\mathbf{E}^I(\mathbf{r}, t) \equiv |E^I\rangle = \sqrt{P_l} \sqrt{2/\epsilon_0} X_l(\mathbf{r}) e^{i k \hat{s}_l \cdot (\mathbf{r} - \mathbf{r}_l)} e^{i \Phi(i)} e^{-i \omega t} \hat{e}_l \quad (7.10)$$

In the matrix elements we express the laser beam in the XY-polarization bases in the I-system: $|E^I\rangle \equiv |E^I\rangle(\mathbf{r}) = |l^I_x\rangle \mathcal{E}^I_{lx} + |l^I_y\rangle \mathcal{E}^I_{ly}$. That base modes are

$$\begin{array}{ll}
|l^I_x\rangle & \equiv \hat{E}^I_{xl}(\mathbf{r}) = \ldots \sqrt{2/\epsilon_0} X_l(\mathbf{r}) e^{i k \hat{x} \cdot (\mathbf{r} - \mathbf{r}_l)} \hat{x} \\
|l^I_y\rangle & \equiv \hat{E}^I_{yl}(\mathbf{r}) = \ldots \sqrt{2/\epsilon_0} X_l(\mathbf{r}) e^{i k \hat{y} \cdot (\mathbf{r} - \mathbf{r}_l)} \hat{y} \\
\end{array} \quad (7.11)$$

Recall from Sec. 5.2 that the base modes are a kind of unit vectors, normalized such that $\langle l|l\rangle = 1$. Therefore the factor $\sqrt{P}$ is missing. The phase of a base mode is irrelevant but for convenience we include the reference point $\mathbf{r}_l$. The corresponding base modes in the O-system (observation modes) are

$$\begin{array}{ll}
|o_x\rangle & \equiv \hat{E}^O_x(\mathbf{r}) = \ldots \sqrt{2/\epsilon_0} X_o(\mathbf{r}) e^{i k \hat{x} \cdot (\mathbf{r} - \mathbf{r}_o)} \hat{x} \\
|o_y\rangle & \equiv \hat{E}^O_y(\mathbf{r}) = \ldots \sqrt{2/\epsilon_0} X_o(\mathbf{r}) e^{i k \hat{y} \cdot (\mathbf{r} - \mathbf{r}_o)} \hat{y} \\
\end{array} \quad (7.12)$$

The Cartesian unit vectors $\hat{x}, \hat{y}_I, \text{and} \hat{y}_O$ are defined in Fig. 7.1. We insert these base modes in place of the fields in Eq. 7.6 and calculate the elements $T_{ol}$ of the $\mathbf{T}^L_{lo}$-matrix. Denoting with $\hat{e}_l$ and $\hat{e}_o$ the involved pairs of polarization vectors, we write

$$T_{ol} = \frac{i k}{2} \frac{1}{\epsilon_o} \int\int_V X_o(\mathbf{r}) X_l(\mathbf{r}^\prime) e^{i k \hat{s}_l \cdot (\mathbf{r} - \mathbf{r}_l)} e^{-i k \hat{s}_o \cdot (\mathbf{r} - \mathbf{r}_o)} [\hat{e}_o \cdot \hat{e}_o \mathbf{X}(\mathbf{r}, \mathbf{r}^\prime) \cdot \hat{e}_l] \, d^3\mathbf{r} \, d^3\mathbf{r}^\prime. \quad (7.13)$$
Now comes a first approximation. We assume that the size of the scattering particle is much smaller than the beam radii $a_l$ and $a_o$ so that $X_o(r) X_i(r')$ are approximately constant across the scattering particle. Thus, we take the profiles out of the integration:

$$T_{ol} = i \frac{k^2}{2 \epsilon_o} X_o(r_m) X_i(r_m) \int \int_V e^{i k \hat{s}_i \cdot (r' - r)} e^{-i k \hat{s}_o \cdot (r - r')} [\hat{e}_o \cdot \epsilon_o \mathbf{X}(r, r') \cdot \hat{e}_l] d^3 r' d^3 r. \quad (7.14)$$

Now we transform the integration variables from $r$ and $r'$ to $u = r - r_m$ and $u' = r' - r_m$ and take the part of the phasor that does not depend on $u$ or $u'$ out of the integration. The remaining integral defines an auxiliary quantity $A_m$ that we call multipole polarizability tensor:

$$A_m = \int \int_V e^{i k \hat{s}_i \cdot u} e^{-i k \hat{s}_o \cdot u} \epsilon_o \mathbf{X}(u, u') d^3 u' d^3 u. \quad (7.15)$$

This $A_m$ is a generalized version of the dipole polarizability tensor introduced in Eq. 2.41. We also introduce the dimensionless normalized version $A = A_m / \alpha$, where $\alpha = \text{Trace}(A_m) / 3$ is a generalized polarizability. Finally, we re-arrange the phasor $\exp[i k \hat{s}_i \cdot (r_m - r_i)] \exp[-i k \hat{s}_o \cdot (r_m - r_f)]$, so that the matrix element can be written as

$$T_{ol} = X_o(r_m) X_i(r_m) e^{i q \cdot r_m} e^{i k L_{if}} \frac{i k \alpha}{2 \epsilon_o} [\hat{e}_o \cdot A \cdot \hat{e}_l] \quad (7.16)$$

In $k L_{if}$ we included the phase contribution that does not depend on the position $r_m$ of the scatterer relative to the focus $r_f$. As shown in Fig. 7.2, the quantity $k L_{if}$ represent the optical path length from $r_i$ to $r_f$. One easily finds that $k L_{if} = k \hat{s}_i \cdot (r_f - r_i)$. Note that we regard the medium as isotropic, both polarizations propagate with the same phase speed and in the same beam. This is a quite serious simplifying approximation, considering the fact that biological tissues can be birefringent. The second phasor $\exp(i q \cdot r_m)$ measures the particle position. The vector $q$ is the so called scattering vector that is defined as

$$q = k_i - k_o = k (\hat{s}_i - \hat{s}_o) \quad (7.17)$$

The magnitude of the scattering vector decreases with the scattering angle $\theta$, since $q^2 = 2k^2[1 - \cos(\theta)]$. Note in Fig 7.2 the fringes in the illuminated and observed region (IOR); their wave fronts are perpendicular to $q$ and their wavelength is $\lambda_q = 2\pi / q$. Note also that $\hbar q = \Delta \hbar k$ is the momentum transferred from the photon to the scattering molecule. By using in the filter $|o\rangle$ the same wave number $k = \omega / c$ as in the input state $|E^i\rangle$, we are restricting the discussion to elastic or quasi-elastic scattering: no change
of photon energy $\hbar \omega$ in the scattering process, only the direction of the momentum $\hbar \mathbf{k}$ changes. Note that in $\exp(i \mathbf{q} \cdot \mathbf{r}_m)$ the position $\mathbf{r}_m$ is coded much more accurately than in the smooth profiles $X_o(\mathbf{r}_m), X_i(\mathbf{r}_m)$. Therefore, the scattering vector $\mathbf{q}$ plays a central role in the analysis of the structure of a matter by scattering techniques.

### 7.4 Amplitude scattering matrix

Equation 7.16 is quite general, the only assumption being that the scattering particle is small, so that the illumination and observation modes can be viewed as plane waves. Unfortunately, however, direct calculation of $\mathbf{A}_m$ is a formidable task, too difficult for the present aims and present skills. Therefore we cast Eq. 7.8 in a canonical form allowing to exploit the tremendous amount of theoretical and experimental work that has been done on arbitrary scattering particles by various techniques. Upon inserting the matrix elements from Eq. 7.16 into Eq. 7.8 we obtain after some re-arrangements:

$$\begin{pmatrix} \mathcal{E}^O_x \\ \mathcal{E}^O_y \end{pmatrix} = \mathcal{B}(\mathbf{r}_m) \left\{ \frac{i k^3}{4 \pi \epsilon_o} \begin{pmatrix} \hat{x} \cdot \mathbf{A} \cdot \hat{x} & \hat{x} \cdot \mathbf{A} \cdot \hat{y}_1 \\ \hat{y}_O \cdot \mathbf{A} \cdot \hat{x} & \hat{y}_O \cdot \mathbf{A} \cdot \hat{y}_1 \end{pmatrix} \right\} \begin{pmatrix} \mathcal{E}^I_x \\ \mathcal{E}^I_y \end{pmatrix}$$

(7.18)

where

$$\mathcal{B}(\mathbf{r}_m) = X_i(\mathbf{r}_m) X_o(\mathbf{r}_m) \frac{2 \pi}{k^2} \epsilon^I q \mathbf{m} e^{i k L I}$$

(7.19)

The quantity in curly brackets is the classical amplitude scattering matrix $\mathbf{S}$, see, e.g., Sec. 3.2. in [24]. Keep in mind that an $\mathbf{S}$-matrix is always defined in the I-system! The present notation is somewhat unusual. For convenience, we decomposed the $\mathbf{S}$-matrix into the pre-factor that is proportional to the generalized polarizability $\alpha$, and the polarization matrix $\mathbf{\Pi}$ that concerns only the polarization transfer:

$$\mathbf{S} = \frac{i k^3}{4 \pi \epsilon_o} \begin{pmatrix} \hat{x} \cdot \mathbf{A} \cdot \hat{x} & \hat{x} \cdot \mathbf{A} \cdot \hat{y}_1 \\ \hat{y}_O \cdot \mathbf{A} \cdot \hat{x} & \hat{y}_O \cdot \mathbf{A} \cdot \hat{y}_1 \end{pmatrix} \quad \mathbf{\Pi} = \begin{pmatrix} \hat{x} \cdot \mathbf{A} \cdot \hat{x} & \hat{x} \cdot \mathbf{A} \cdot \hat{y}_1 \\ \hat{y}_O \cdot \mathbf{A} \cdot \hat{x} & \hat{y}_O \cdot \mathbf{A} \cdot \hat{y}_1 \end{pmatrix}$$

(7.20)

The standard notation is

$$\mathbf{S} = \begin{pmatrix} S_{xx} & S_{xy} \\ S_{yx} & S_{yy} \end{pmatrix} = \begin{pmatrix} S_1 & S_4 \\ S_3 & S_2 \end{pmatrix}$$

(7.21)

The numbering of the elements $S_1$-$4$ corresponds to the notation in [38]. Working in the $xy$-linear base, the four matrix elements concern the following polarization transfers:

$$\begin{align*}
S_1 : \hat{e}_I \perp = \hat{x} \rightarrow \hat{x} = \hat{e}_O \perp & \quad S_4 : \hat{e}_I \parallel = \hat{y}_I \rightarrow \hat{x} = \hat{e}_O \perp \\
S_3 : \hat{e}_I \perp = \hat{x} \rightarrow \hat{y}_O = \hat{e}_O \parallel & \quad S_2 : \hat{e}_I \parallel = \hat{y}_I \rightarrow \hat{y}_O = \hat{e}_O \parallel
\end{align*}$$

(7.22)

The canonical form $|E^o\rangle = \mathcal{B}(\mathbf{r}_m) \mathbf{S} |E^I\rangle$ is applicable to the scattering of a quasi-plane wave by an arbitrary particle of arbitrary size $a$ such that $a \ll a_l$. The dimensionless factor $\mathcal{B}(\mathbf{r}_m)$ corresponds to the factor $\exp(i k r)/k r$ of the traditional far-field formulation of the scattering problem. The far-field result can be easily derived using the present mode projection approach, one only has to replace the beam filter (that consists of a lens and fiber or pinhole, Fig. 5.1) with a dipole antenna far from the scatterer.

### 7.5 Scattering signal, cross-sections and phase function

To understand the significance of the $\mathbf{S}$-matrix, we consider the measurable quantity, namely the received power $P_o$. Since we removed the polarizer from the detection
channel, the received power is the sum of two contributions from the two orthogonal modes: \( P_{os} = |E^{os}|^2 = |\mathcal{E}_x^{o}|^2 + |\mathcal{E}_y^{o}|^2 \), where

\[
|\mathcal{E}_x^{o}|^2 = |T_{xx}E_x^f + T_{xy}E_y^f|^2 = |\mathcal{B}(r_m)|^2 |S_{xx}\mathcal{E}_x^l + S_{xy}\mathcal{E}_y^l|^2 \\
|\mathcal{E}_y^{o}|^2 = |T_{yz}E_x^f + T_{yy}E_y^f|^2 = |\mathcal{B}(r_m)|^2 |S_{yz}\mathcal{E}_x^l + S_{yy}\mathcal{E}_y^l|^2
\] (7.23)

In the concise Jones vector notation this reads

\[
P_{os} = \left| E_i^l \right|^2 \left| T_{lo}^l \right|^2 \left| E_o^l \right|^2 = |\mathcal{B}(r_m)|^2 \left| E_i^l \right|^2 \left| S^l S^l \right| E_o^l^\dagger
\] (7.24)

Most often we do not care about the absolute power received, but for the ratio of input and output, i.e., in the probability \( P_{os}(\theta, \phi) \) that a photon traveling in the laser beam is scattered into the observation beam aligned along \( \hat{s}_o \equiv (\theta, \phi) \):

\[
P_{os}(\theta, \phi) = \frac{P_{os}}{P_i} = \left| \mathcal{B}(r_m) \right|^2 \frac{\left| E_i^l \right|^2 \left| S^l S^l \right| E_o^l^\dagger}{\left| E_i^l \right|^2} = \left| \mathcal{B}(r_m) \right|^2 \frac{\left| E_i^l \right|^2 \left| R^l_{21}^{-1}(\phi) S^l S \right| R^l_{21}(\phi) E_o^l^\dagger}{\left| E_i^l \right|^2}
\] (7.25)

where, of course \( \left| E_i^l \right|^2 = \left| E_i^l \right|^2 = P_i = \left| \mathcal{E}_x^l \right|^2 + \left| \mathcal{E}_y^l \right|^2 \). This looks like a trivial re-arrangement, but one should realize that the complex amplitudes can be interpreted as \( \mathcal{E}_x^l = \sqrt{P_i} e^{i\theta} e_{xl}^l \) and \( \mathcal{E}_y^l = \sqrt{P_i} e^{i\phi} e_{yl}^l \). Here \( e_{xl}^l \) and \( e_{yl}^l \) are the components of a complex unit polarization vector in the I-system, such that \( |e_{xl}^l|^2 + |e_{yl}^l|^2 = 1 \). Thus, the fraction \( \left| E_i^l \right|^2 \left| S^l S^l \right| E_o^l^\dagger / \left| E_i^l \right|^2 \) on the right hand side of Eq. 7.25 depends only on the polarization but not on the power and phase of the laser beam. In order to establish a link with the common radiometric language, we note that the pre-factor \( \left| \mathcal{B}(r_m) \right|^2 \) (recall \( \mathcal{B} \) from Eq. 7.19) can be expressed as follows

\[
\left| \mathcal{B}(r_m) \right|^2 = X_i(r_m)Y_o(r_m)\Omega_o \frac{1}{k^2}
\] (7.26)

The quantities \( \Omega_o \) and \( Y_o \) have been defined in Sec. 2.3.2 in the context of Eq. 2.40: \( \Omega_o = 4\pi/k^2a_o^2 \) is the effective solid angle accepted by the receiver, \( a_o \) is the waist radius of the observation mode. The dimensionless profile \( Y_o(r_m) = \pi a_o^2 X_i^2(r_m) \) quantifies the alignment of the receiver; one can set \( Y(r_m) = 1 \) if the scatterer is exactly on the beam axis. Further we recall from Sec. 2.3.2 the expression \( I(r_m) = P_i X_i^2(r_m) \) for the local irradiance [W m\(^{-2}\)] of the laser beam at the position of the scatterer. Using these definitions, we can re-arrange Eq. 7.25 into a form that defines the quantity \( \partial\sigma \), known as the differential scattering cross-section:

\[
\partial\sigma_s(\hat{s}_o|\hat{s}_i, \hat{e}_l) = \frac{1}{Y_o(r_m)} \frac{P_{os}}{\Omega_o} = \frac{1}{k^2} \frac{\left| E_i^l \right|^2 \left| S^l S^l \right| E_o^l^\dagger}{\left| E_i^l \right|^2}
\] (7.27)

"Physically, \( \partial\sigma \) specifies the angular distribution of the scattered light: the amount of light (for unit incident irradiance) scattered into a unit angle about a given direction." The factor \( 1/Y_o(r_m) \) corrects for the misalignment of the receiver. For future use, we note \( \partial\sigma \) Eq. 7.27 in the following form

\[
\partial\sigma_s(\hat{s}_o|\hat{s}_i, \hat{e}_l) = \frac{ik^4}{6\pi} \left| e_o \right|^2 \left| \alpha \right|^2 \frac{3}{8\pi} \Pi \quad \text{where} \quad \Pi = \left\langle e_i^l | \Pi^l \Pi | e_l^l \right\rangle
\] (7.28)

where \( \Pi \) is the polarization transfer matrix, Eq. 7.20. The right hand side defines the polarization factor \( \Pi \). Note that in general \( \partial\sigma_s(\hat{s}_o|\hat{s}_i, \hat{e}_l) \) depends on the orientation of the scattering particle with respect to the propagation direction and polarization of
the incident field. This should be kept in mind when subsequently using $\partial \sigma_t(\theta, \phi)$ as an abbreviation for $\partial \sigma_t(\hat{s}_o|\hat{s}_i, \hat{e}_i) = \partial \sigma_t(\theta, \phi|\hat{s}_i, \hat{e}_i)$. By integrating $\partial \sigma(\theta, \phi)$ over the $4\pi$ solid angle one obtains the total scattering cross-section $\sigma_t$:

$$\sigma_t(\hat{s}_i, \hat{e}_i) = \int_{\theta=0}^{\pi} \int_{\phi=0}^{2\pi} \partial \sigma(\theta, \phi) \sin(\theta) \, d\phi \, d\theta$$  \hspace{1cm} (7.29)$$

In an elementary interpretation, $\sigma_t$ can be viewed as the effective area of the molecules exposed to a photon stream. (Thus, a scattering system that consists of identical particles at a mean concentration $\bar{c}$ will exhibit a scattering coefficient $\mu_s = \sigma_t \bar{c}$; see Sec. 8.2.1.) To illustrate the concept of cross-sections we cast Eq. 7.27 in a form, which can be interpreted in probabilistic terms. The probability $P_{os} = P_{os}/P_t$ that a photon is scattered from the laser beam into the observation beam is

$$P_{os} = \Omega_o \frac{\sigma_t}{\pi a^2} p(\hat{s}_o|\hat{s}_i, \hat{e}_i) Y_o(r_m) Y_t(r_m)$$  \hspace{1cm} (7.30)$$

The ratio $\sigma_t/\pi a^2$ is the geometrical probability that a photon traveling in a beam of an effective cross-section $\pi a^2$ is scattered by a target of cross-section $\sigma_t$. The factor $Y_o(r_m)Y_t(r_m)$ accounts for misalignment of illumination and observation. The product $\Omega_o p(\hat{s}_o|\hat{s}_i, \hat{e}_i)$ is the probability that a scattered photon flies into the solid angle $\Omega_o$ around the direction $\hat{s}_o$; the corresponding angular probability density is

$$p(\hat{s}_o|\hat{s}_i, \hat{e}_i) = p(\theta, \phi|\hat{s}_i, \hat{e}_i) = \frac{\partial \sigma(\hat{s}_o|\hat{s}_i, \hat{e}_i)}{\sigma_t(\hat{s}_i, \hat{e}_i)}$$  \hspace{1cm} (7.31)$$

Note that the probability density $p(\theta, \phi|\hat{s}_i, \hat{e}_i)$ is per definition normalized such that $\int p(\theta, \phi|\hat{s}_i, \hat{e}_i) \sin(\theta) \, d\theta \, d\phi = 1$. Note also that $p(\theta, \phi|\hat{s}_i, \hat{e}_i)$ is a generalized version of the “phase function” that was introduced in Chapter 3. Consequently we also define the asymmetry parameter

$$g = \langle \hat{s}_i \cdot \hat{s}_o \rangle = \langle \cos(\theta) \rangle = \int_{\theta=0}^{\pi} \int_{\phi=0}^{2\pi} p(\theta, \phi) \cos(\theta) \sin(\theta) \, d\phi \, d\theta$$  \hspace{1cm} (7.32)$$

This $g$ is more often called “anisotropy factor”, but we prefer “asymmetry parameter” in order to avoid confusion with anisotropy of shape and dielectric response of the particle. Even in isotropic systems, the phase function is an asymmetric function of the scattering angle $\theta$, because of the $\theta$-dependence of the scattering vector $q$.

In the context of radiative transfer Eq. 7.30 is often expressed in radiometric terms, namely in terms of the radiant intensity $I_\Omega$ and the irradiance $I$:

$$I_\Omega(\hat{s}_o) = \partial \sigma_s(\hat{s}_o|\hat{s}_i, \hat{e}_i) I(r_m) Y_o(r_m) = p(\hat{s}_o|\hat{s}_i, \hat{e}_i) \sigma_t(\hat{s}_i, \hat{e}_i) I(r_m) Y_o(r_m)$$  \hspace{1cm} (7.33)$$

Here $I_\Omega(\hat{s}_o) = P_{os}/\Omega_o$ is the radiant intensity [W/sr] of the beam wave selected from the scattered field to propagate in the direction $\hat{s}_o$ and $I(r_m) = P Y(r_m)/\pi a^2$ is the irradiance [W/m²] in the laser beam at the position $r_m$ of the scatterer; recall Eq. 2.40. The alignment factor $Y_o(r_m)$ is usually neglected.

For future use in a Monte Carlo algorithm we note that the “phase function” can be directly expressed in terms of the normalized S-matrix:

$$\hat{S} = \frac{1}{\sqrt{K^2 \sigma_t}} S,$$  \hspace{1cm} (7.34)$$

so that

$$p(\hat{s}_o|\hat{s}_i, \hat{e}_i) = \frac{\langle E'_l | \hat{S}^* \hat{S} | E'_l \rangle}{\langle E'_l | E'_l \rangle} = \frac{\langle E'_l | R_2^{-1} \hat{S}^* \hat{S} R_2 | E'_l \rangle}{\langle E'_l | E'_l \rangle}.$$  \hspace{1cm} (7.35)$$
A considerable simplification arises for particles that are isotropic and non-chiral, both, in shape and in dielectric response. Then the off-diagonal elements $S_3$ and $S_4$ of the S-matrix are zero (Sec. 13.6. in [24]), and the phase function can be written as

$$p(\theta, \phi) = |e^f_{3t}|^2 |\hat{S}_3|^2 + |e^f_{gt}|^2 |\hat{S}_2|^2$$

$$= \frac{|\hat{S}_3|^2 + |\hat{S}_2|^2}{2} \left\{ \cos(2\phi) \left[ |e^f_{3t}|^2 - |e^f_{gt}|^2 \right] + 2 \sin(2\phi) \Re(e^f_{3t} e^f_{gt}) \right\},$$

where $e^f_{3t}$, $e^f_{gt}$ and $e^f_{1t}$ are the components of the polarization vector $\hat{e}_i$ in the I-system and L-system, respectively.

### 7.6 Rayleigh scatterer

In the present chapter we consider a truly elementary scatterer, namely a particle that is small enough with respect to the light wavelength so that it can be regarded as a point. Thus, the oscillating charge distribution excited by the illuminating field is an infinitesimal dipole oscillator $d$ localized at $\delta(r - r_m)$:

$$d = A_d \cdot E'(r_m) e^{-i\omega t}.$$  (7.37)

Here $A_d$ is the dipole polarizability tensor (recall Eq. 2.41) and $E'(r_m)$ is the illuminating field from Eq. 7.10. The magnitude of $d$ follows the field oscillations, but the direction $\hat{d}$ does not necessarily coincide with the impinging polarization. For example, in the case of uniaxial anisotropy with anisotropy axis $\hat{a}$ one may write explicitly

$$d \sim A_d \cdot \hat{e}_i = \alpha_i \hat{e}_i + \Delta \alpha \ \hat{a} \cdot \hat{a} \cdot \hat{e}_i = \alpha \hat{A} \cdot \hat{e}_i$$  (7.38)

where $\alpha = \text{Trace}(A_d)/3 = \alpha_i + \Delta \alpha/3$. In the absence of anisotropy, the normalized matrix $\hat{A}$ reduces to the unity matrix $I$ (1 on diagonal, 0 otherwise). The oscillating dipole radiates the scattered field $|E|^\alpha$ (recall Eq. 2.43) whose source is the current density distribution

$$j(r, t) = -i \omega \delta(r - r_m) d(r_m) = -i \omega \delta(r - r_m) A_d \cdot E'(r, t).$$  (7.39)

This $j(r, t)$ is just a simple special version of the general current density from Eq. 7.5, with $c_s X(r-r')$ being replaced by $\delta(r-r_m) \delta(r-r') A_d$. Thus, starting from Eq. 7.13, one verifies that all to be done is to replace the multipole polarizability tensor $A_m$ with the dipole tensor $A_d$. With the simple dipole, it is quite easy to get an explicit expression for the total scattering cross-section $\sigma_d$. The power $P_s$ radiated in the scattered field is obtained upon inserting the dipole moment $d$ from Eq. 7.37 into Eq. 2.44 and just a little work is needed to find

$$\sigma_d(\hat{e}_i) = \frac{k^4 |\alpha|^2}{6\pi} \left( \hat{e}_i^\dagger \cdot A^\dagger \cdot \hat{e}_i \right).$$  (7.40)

Notice in $\sigma_d$ the $k^4$-dependence. The scattering intensity increases with $1/\lambda^4$, which is characteristic for the so-called Rayleigh scatterers. Notice also that $\sigma_d(\hat{e}_i)$ depends in general on the orientation of an anisotropic molecule with respect to the impinging polarization $\hat{e}_i$, as expressed by the anisotropy factor $\langle \hat{e}_i^\dagger \cdot A^\dagger \cdot \hat{e}_i \rangle$. For example, in the uniaxial case one obtains

$$\langle \hat{e}_i^\dagger \cdot A^\dagger \cdot \hat{e}_i \rangle = \frac{\alpha_i}{\alpha} \left\{ 1 + \gamma (1 + \gamma) |\hat{a} \cdot \hat{e}_i|^2 \right\}, \quad \text{where} \quad \gamma = \Delta \alpha/\alpha_i$$  (7.41)
When calculating the cross-section of the elements of the polarization transfer matrix $\Pi$ (recall Eq. 7.20; e.g. $\hat{x} \cdot A \cdot \hat{y}$), one should not forget that all involved quantities must be in the same coordinate system. Since the tensor $A$ is a property of the scattering sample, one would transform the vectors into the lab-frame. Such transformations make the anisotropic scattering a tedious algebraic exercise, best done by a computer. A thorough discussion of Rayleigh scattering, including explicit expressions for polarizability tensors of various types of Rayleigh particles, can be found in Chapter 5 in [24].

Here we content ourselves with a compilation of the explicit results for an isotropic dipole scatterer. We replace in all formulas the dimensionless polarizability tensor $A$ by the unity matrix $I$. This implies in particular that in Eq. 7.40 for the total cross-section $\sigma_d$ we set $(\hat{e}_l^I \cdot A^I \cdot \hat{e}_l^I) = \hat{e}_l^I \cdot \hat{e}_l^I = 1$. The elements of the isotropic polarization transfer matrix $\Pi$ can be more or less looked up in Fig. 7.3, so that we just write

$$
\Pi_d = \begin{pmatrix} 1 & 0 \\ 0 & \cos(\theta) \end{pmatrix} \quad \mathbf{S}_d = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \begin{pmatrix} \cos(\theta) \end{pmatrix}.
$$

(7.42)

Such diagonal structure of the $S$-matrix, with $|S2/S1|^2 = \cos^2(\theta)$, is characteristic for single scattering from an isotropic system that fulfills the Born approximation. We also recall the normalized version of the $S$-matrix

$$
\hat{S}_d = \frac{1}{\sqrt{k^2\sigma_d}} \mathbf{S}_d = \sqrt{\frac{3}{8\pi}} \begin{pmatrix} 1 \\ 0 \end{pmatrix} \begin{pmatrix} \cos(\theta) \end{pmatrix}.
$$

(7.43)

from which we calculate the phase function of an isotropic dipole according to Eq. 7.36:

$$
p_d(\theta, \phi|\hat{e}_I) = \frac{3}{8\pi} \Pi_d(\theta, \phi|\hat{e}_I),
$$

(7.44)

where $\Pi_d(\theta, \phi|\hat{e}_I)$ is the dipole polarization factor defined in Eq. 7.28:

$$
\Pi_d(\theta, \phi|\hat{e}_I) = \frac{|e_{xl}^I|^2 + |e_{yl}^I|^2 \cos^2(\theta)}{1 + \cos^2(\theta)} + \frac{\sin^2(\theta)}{2} \left\{ \cos(2\phi) \left[ |e_{xl}^I|^2 - |e_{yl}^I|^2 \right] + 2 \sin(2\phi) \Re(e_{xl}^I e_{yl}^I) \right\}.
$$

(7.45)

Here $e_{xl}^I, e_{yl}^I$ and $e_{xl}^L, e_{yl}^L$ are the components of the polarization vector $\hat{e}_I$ in the I-system and L-system, respectively. One can easily verify that $\int\int p(\theta, \phi) \sin(\theta) d\theta d\phi = 1$. We
also note a couple of special cases. With a vertically polarized laser, so that $\hat{e}_1 = \hat{x}_L$, one obtains $\Pi_{\vartheta}(\theta, \phi|\hat{x}_1) = \cos^2(\phi) + \sin^2(\phi) \cos^2(\theta)$. With a circularly polarized laser, only the first term in Eq. 7.45 survives, and the same also applies on average for the so-called “natural” light, i.e. for randomly fluctuating polarization with uniform distribution of the polarization vectors $\hat{e}_1$. 
Chapter 8

Scattering from interacting Rayleigh particles.

The most simple heterogeneous optical medium is a gas consisting of identical isotropic molecules. With the dilute gas we perform now a scattering experiment of the kind as discussed in chapter 7.1 and illustrated in Fig. 8.1. The only difference is that now the illuminated and observed region IOR formed by the profile overlap $X_i(r)X_f(r)$ contains many scatterers. The gas molecules are Rayleigh scatterers: when illuminated with a laser beam, they respond to $\mathbf{E}^l$ with dipole oscillations according to $\mathbf{X}_m = \alpha \mathbf{E}^l$. These oscillators are sources of molecular dipole fields $\mathbf{E}^m$ (Eqs. 2.43 and 2.42), which in turn contribute to the excitation of neighbor molecules. However, since the polarizability $\alpha$ is small and a gas is dilute, we neglect these secondary fields assuming that $\mathbf{E}^l \gg \mathbf{E}^m$. In other words, we neglect all kinds of multiple scattering and assume that the molecular dipoles are driven by the unperturbed laser field $\mathbf{E}^l$, as it would exist in the absence of the scatterers. This is called first order Born approximation.

In the old Maxwell’s days, before the existence of molecules was established, one would understand this gas as a dielectric continuum with a susceptibility $\chi = \gamma c_{\text{mol}}$, where $c_{\text{mol}}$ is the molar concentration (in mol/l) and $\gamma$ is the molar specific susceptibility. Then one would use the constitutive relation Eq. 2.15: $\mathbf{P} = \epsilon_o \chi \mathbf{I} \cdot \mathbf{E}^l$, where $\mathbf{P}$ is the density of induced dipole moments. (The unity matrix $\mathbf{I}$ indicates explicitly that we restrict the discussion to isotropic response.) We are convinced of the existence of molecules, but we can still use the continuum concept for a short hand notation:

$$\mathbf{P}(\mathbf{r},t) = \epsilon_o \chi(\mathbf{r},t) \mathbf{I} \cdot \mathbf{E}^l(\mathbf{r}) \quad \text{where} \quad \chi(\mathbf{r},t) = \frac{\alpha}{\epsilon_o} c_m(\mathbf{r},t) \quad (8.1)$$

Here $c_m(\mathbf{r},t)$ is the microscopic molecular concentration describing the distribution of the gas molecules. On the microscopic scale, the dilute gas is a highly inhomogeneous medium; $c_m(\mathbf{r},t)$ consists of $\delta$ spikes, a $\delta$-function for each molecule $m$ at its momentaneous position $\mathbf{r}_m(t)$:

$$c_m(\mathbf{r},t) = \sum_m \delta[\mathbf{r} - \mathbf{r}_m(t)] \quad (8.2)$$

Because of thermal motions of the gas molecules, the positions of the $\delta$-spikes keep changing with time; $c_m(\mathbf{r},t)$ is a random and fluctuating distribution. The fluctuations of $c_m(\mathbf{r},t)$ are much slower than light oscillations $\exp(-i\omega t)$, which allows to write the source of the scattered field, namely the current density distribution $\mathbf{j}(\mathbf{r},t) = \partial \mathbf{P} / \partial t$ as

$$\mathbf{j}(\mathbf{r},t) = -i\omega c_m(\mathbf{r},t) \alpha \mathbf{I} \cdot \mathbf{E}^l(\mathbf{r}) = -i\omega \epsilon_o \chi(\mathbf{r},t) \mathbf{I} \cdot \mathbf{E}^l(\mathbf{r}). \quad (8.3)$$

This $\mathbf{j}(\mathbf{r},t)$ is a special version of the general current density from Eq. 7.5, with $\mathbf{X}(\mathbf{r} - \mathbf{r}')$ being replaced by $c_m(\mathbf{r}) \alpha \mathbf{I} \delta(\mathbf{r} - \mathbf{r}')$. Thus, starting from Eq. 7.13, one finds that each
The elements $T_{ol}$ of the $T^{l}_{lo}$ matrix has the form

$$T_{ol}(t) = e^{ikL_{o}t} \frac{i k}{2} \alpha \int X_{o}(r)X_{l}(r) c_{m}(r,t) e^{i q \cdot r} d^{3}r (\hat{e}_{l} \cdot \hat{e}_{o}^{*}) \quad (8.4)$$

$$= e^{ikL_{o}t} \frac{i k}{2} \int X_{o}(r)X_{l}(r) \chi(r,t) e^{i q \cdot r} d^{3}r (\hat{e}_{l} \cdot \hat{e}_{o}^{*}) \quad (8.5)$$

The second version is to keep in mind the continuum model, to which we shall return later. The Fourier integral in Eq. 8.4 expresses the interference of the random contributions of the individual scatterers to the scattering amplitude. Neglecting the broad IOR-profile $X_{l}(r)X_{o}(r)$, the amplitude is proportional to the $q$-Fourier component of the particle distribution $c_{m}(r)$. This is the fundament of scattering experiments aimed on the elucidation of the structure of the matter. By choosing a certain $q(\theta)$, one chooses the probing length scale, i.e., the wavelength $\lambda = \frac{2 \pi}{q}$ of a spatial harmonic in $c(r)$. Recall that $q^{2} = k^{2} |\hat{s}_{o} - \hat{s}_{l}|^{2} = 2k^{2} [1 - \cos(\theta)]$. Thus, with light scattering the shortest probing scale is $\frac{1}{2k} = \frac{\lambda}{4\pi}$, i.e. to roughly 50nm in visible range.

The wavelength $\lambda$ corresponds to the spacing of the fringes that are visible in the IOR in Fig. 7.2. As the particles move randomly through the IOR, the scattering amplitude fluctuates and this also results in the fluctuations of the signal $P_{os}(t)$ (recall the definition in Eq. 7.23). The measured quantity is the time averaged signal:

$$\langle P_{os}(t) \rangle = \langle |T_{xx}(t)|^{2} \rangle |\mathcal{E}_{lx}|^{2} + \langle |T_{yy}(t)|^{2} \rangle |\mathcal{E}_{ly}|^{2} \quad (8.6)$$

The mean square matrix elements are readily evaluated to

$$\langle |T_{ol}(t)|^{2} \rangle = \frac{\sigma_{d}}{\pi a_{l}^{2}} \Omega_{o} \frac{3}{8\pi} |\hat{e}_{l} \cdot \hat{e}_{o}^{*}|^{2} S(q) \quad (8.7)$$

where

$$S(q) = \int \int Y_{ol}(r) \langle c_{m}(r)c_{m}(r') \rangle e^{i q \cdot (r-r')} Y_{ol}(r') d^{3}r d^{3}r' \quad Y_{ol} = X_{l}X_{o}a_{l}a_{o} \quad (8.8)$$

In the factor $(3/8\pi)|\hat{e}_{l} \cdot \hat{e}_{o}^{*}|^{2}$ we recognize the squared elements of the normalized S-matrix $\hat{S}^{d}$, Eq. 7.43. Thus, recalling also Eq. 7.36 we find that the normalized scattering signal $P_{os} = P_{os}/P_{l}$ can be written as

$$\langle P_{os} \rangle = \Omega_{o} \frac{1}{\pi a_{l}^{2}} \frac{k^{4}}{6\pi} \frac{\alpha_{d}}{\epsilon_{o}} \left| p_{d}(\hat{s}_{o} |\hat{s}_{l}, \hat{e}_{l}) \right|^{2} \sigma_{d} \quad (8.9)$$
where \((k^4/6\pi)|\alpha_d/\varepsilon_0|^2 = \sigma_d\) is the total differential cross-section of an elementary dipole and \(p_d(\theta, \phi)\) is the dipole phase function from Eqs. 7.44 and 7.45. When comparing Eq. 8.9 with Eq. 7.30 for scattering from an arbitrary particle, one notices only one difference: the alignment factor \(Y_o(\mathbf{r}_m)Y_l(\mathbf{r}_m)\) is replaced by the integral \(\mathcal{S}(\mathbf{q})\).

### 8.1 Scattering and fluctuations

Notice that in the integral in Eq. 8.8 a new quantity has appeared, namely the *spatial correlation function* \(g(\mathbf{r}, \mathbf{r}') = \langle c_m(\mathbf{r})c_m(\mathbf{r}')\rangle\) of the distribution \(c_m(\mathbf{r})\). Recalling Eq. 8.2 we see that \(g(\mathbf{r}, \mathbf{r}')\) is the average of a double sum of \(\delta\)-functions:

\[
\langle c_m(\mathbf{r})c_m(\mathbf{r}')\rangle = \sum_m \sum_n \langle \delta[\mathbf{r} - \mathbf{r}_m(t)][\mathbf{r} - \mathbf{r}_n(t)] \rangle
\] (8.10)

This correlation function is central in the statistical physics of soft condensed matter and likely to play an important role in the understanding of tissue optics. We can’t go too deep into the enormous field of statistical physics, but few hints are necessary to understand the role of fluctuations and interference in light scattering. First, we recall that any random quantity can be written as the sum of a mean and of fluctuations around mean. Thus, we write \(c_m(\mathbf{r}) = \bar{c}(\mathbf{r}) + \bar{c},\) where \(\bar{c}(\mathbf{r}) = \langle c_m(\mathbf{r}) \rangle\) is the macroscopic concentration that is defined as the ensemble of the microscopic distribution from Eq. 8.2. Thus, we decompose the density correlation into two terms, \(g(\mathbf{r}, \mathbf{r}') = \bar{c}^2 + g_\delta(\mathbf{r}, \mathbf{r}')\), where \(g_\delta(\mathbf{r}, \mathbf{r}') = \langle \delta c_m(\mathbf{r})\delta c_m(\mathbf{r}') \rangle\) is the correlation function of density fluctuations. We assume for simplicity that the system is statistically homogeneous, which means that \(g_\delta(\mathbf{r}, \mathbf{r}') = g_\delta(\mathbf{\rho})\) depends only on the difference \(\mathbf{\rho} = \mathbf{r}' - \mathbf{r}\) of the two positions \(\mathbf{r}'\) and \(\mathbf{r}\). (It is sufficient that we require statistical homogeneity within IOR.) Reflecting on the meaning of the different terms in the double sum, Eq. 8.10, one finds that the fluctuation correlation function can be written as the sum of two terms, so that \(g(\mathbf{\rho})\) consists of three terms

\[
\bar{c}^2 + g_\delta(\mathbf{\rho}) = \bar{c}^2 + \bar{c}\delta(\mathbf{\rho}) + \bar{c}^2 h(\mathbf{\rho})
\] (8.11)

The first fluctuation term, the \(\delta\)-correlation peak \(\bar{c}\delta(\mathbf{\rho})\), comes from the \(i = j\) terms of the double sum. This so-called self-correlation peak expresses the obvious fact that each particle is highly correlated with itself. The second fluctuation term \(\bar{c}^2 h(\mathbf{\rho})\), where \(h(\mathbf{\rho})\) is the so-called pair correlation function, expresses the role of pairwise interactions of the particles. This is the contribution from the terms with \(i \neq j\), concerning pairs of two distinct molecules. The positions of the particles are random, but the randomness has its limits, because the molecules do interact and their positions are therefore correlated. This correlation extends over a certain correlation length \(\lambda_c\) that corresponds to the characteristic width of the function \(h(\mathbf{\rho})\).

Inserting \(g(\mathbf{\rho})\) from Eq. 8.11 into Eq. 8.8, one obtains the Fourier integral \(\mathcal{S}(\mathbf{q})\) as the sum of three contributions:

\[
\mathcal{S}(\mathbf{q}) = \mathcal{S}_{fwd} + \mathcal{S}_{inc} + \mathcal{S}_{coh}
\] (8.12)

The first contribution \(\mathcal{S}_{fwd}\) is due to the constant background \(\bar{c}^2\):

\[
\mathcal{S}_{fwd} = \bar{c}^2\left| \int Y_o(\mathbf{r})e^{-i\mathbf{q}\cdot\mathbf{r}}d^3\mathbf{r} \right|^2
\] (8.13)

This integral is a 3D Fourier transform of the IOR profile, whose size is in the order of the beam radius \(a_w\). It follows from properties of Fourier transformation that \(\mathcal{S}_{fwd}\)
will only contribute to forward scattering, i.e. at small \( q^2 < 1/a_w^2 = k^2 \sigma_k^2 \); recall Eq. 2.37. In this forward direction the receiver picks up predominantly the laser beam. Therefore forward scattering is omitted in standard scattering experiments. We can’t do this when propagating light through the tissue. In fact, forward scattering is especially important, because it causes extinction. We shall devote to this tricky topic the chapter 8.2. For now we note an important finding: only the fluctuations \( \delta c(\mathbf{r}, t) \) or \( \delta \chi(\mathbf{r}, t) \) contribute to scattering at non-zero scattering angles.

The second contribution \( S_{inc} \) is due to the self-correlation peak \( \bar{c} \delta(\mathbf{r}) \):

\[
S_{inc} = \bar{c} \int Y_l(\mathbf{r}) Y_o(\mathbf{r}) d^3r = \bar{c} V_{IOR}. \tag{8.14}
\]

Here \( V_{IOR} \) is the effective volume of the IOR and \( \bar{c} V_{IOR} \) is the number of scatterers in the IOR. Each scatterer contributes independently according to its individual differential cross-section \( \sigma_d \), viz. Eq. 8.7. There are no light interference effects and therefore this contribution is called “incoherent”. Interference effects are contained in the third contribution due to \( \bar{c}^2 h(\mathbf{r}) \), which is therefore called the “coherent” term:

\[
S_{coh} = \bar{c}^2 \int \int Y_d(\mathbf{r}) Y_d(\mathbf{r} + \mathbf{\rho}) h(\mathbf{\rho}) e^{-i q \mathbf{\rho} \cdot d^3\mathbf{\rho}}. \tag{8.15}
\]

Here we introduce an additional assumption: the typical correlation length \( \lambda_c \) is much shorter than the size of the IOR, so that \( h(\mathbf{\rho}) \) can be regarded as a \( \delta \)-function when compared with the broad profile \( Y_d(\mathbf{r}) \). Then we can factor the double integration in \( \int Y_d^2(\mathbf{r}) d^3r \cdot \int h(\mathbf{\rho}) e^{-i q \mathbf{\rho} \cdot d^3\mathbf{\rho}} \), and the coherent contribution becomes

\[
S_{coh} = \bar{c} V_{IOR} \mathcal{H}(q) \bar{c} V_C = S_{inc} \mathcal{H}(q) \bar{c} V_C \tag{8.16}
\]

where

\[
\mathcal{H}(q) = \frac{1}{V_C} \int h(\mathbf{\rho}) e^{-i q \mathbf{\rho} \cdot d^3\mathbf{\rho}} \quad V_c = \int h(\mathbf{\rho}) d^3\mathbf{\rho} \tag{8.17}
\]

The function \( \mathcal{H}(q) \) is normalized so that \( \mathcal{H}(0) = 1 \). The normalization constant \( V_c \) can be regarded as a correlation volume that expresses the tendency of the molecules to clustering. Usually, the incoherent and coherent contributions are found combined into \( S = S_{inc} + S_{coh} = \bar{c} V_{IOR} S(q) \), where \( S(q) \) is the so-called structure factor:

\[
S(q) = \frac{1}{\bar{c}} \int \bar{c} e^{i q \mathbf{\rho}} g_3(\mathbf{\rho}) d^3\mathbf{\rho} = 1 + \bar{c} V_c \mathcal{H}(q) \tag{8.18}
\]

The structure factor contains essentially all information about the structure that can be obtained from elastic scattering experiments. In systems with isotropic interactions \( h(\mathbf{\rho}) \) depends only on \( \rho = |\mathbf{\rho}| \) and the structure factor depends only on \( q = |q| = k \sqrt{2 - 2 \cos(\theta)} \). We summarize the achieved. Neglecting forward scattering, the normalized time averaged signal \( \langle P_{os} \rangle_{\theta > 0} \) can be written as

\[
\langle P_{os} \rangle_{\theta > 0} = \Omega_a \frac{1}{\pi a_t^2} \frac{k^4}{6\pi} \left| \frac{\alpha_d}{\epsilon_o} \right|^2 p_d(\theta, \phi^j) S(q) \bar{c} V_{IOR} = \Omega_a \sigma_d p_d(\theta, \phi^j) S(q) \bar{c} V_{IOR} \tag{8.19}
\]

Because of interference, the measured intensity is not the sum of individual intensities. Neglecting the coherent contribution can’t be justified by the vague argument of “randomness”. Actually the mutual arrangement is never quite random because all realistic particles do interact. What matters are the relations between three length scales: the
mean inter-particle distance $d \approx \bar{c}^{1/3}$, their interaction range characterized by the correlation length $\lambda_C \approx V_C^{1/3}$ and the length scale $1/q$ probed in the scattering experiment. Coherent scattering can be neglected if the number of molecules per correlation volume is small, i.e. if $\bar{c}V_c \ll 1$ and/or if $s(q) \ll 1$. The former condition is well fulfilled in the interstellar dust, but not in the biological tissue, where the coherent term dominates in the optically accessible $q$-range.

### 8.2 Forward scattering, extinction and scattering coefficient

Since forward scattering is not forbidden in the propagation through a tissue, we must investigate the case $q = 0$. The experiment is sketched in Fig. 8.2. A laser beam traverses a dilute cloud of interacting Rayleigh particles confined in a slab oriented perpendicularly to the propagation direction. After passing through the slab, the laser beam is expected to be attenuated, because of the scattering. We also expect a slight phase shift with respect to the free propagation, because the cloud of dielectric particles may be regarded as a dielectric medium with mean susceptibility $\bar{\chi} = \bar{c}\alpha/\varepsilon_0$. To analyze the situation, we shall once again employ the projection formalism outlined in section 5 and 7.2. We seek a Jones matrix $T^t$ that characterizes the transmission from the plane at $z = 0$ to the plane $z = d$:

$$
\begin{pmatrix}
E^t_x \\
E^t_y
\end{pmatrix} = T^t
\begin{pmatrix}
E^l_x \\
E^l_y
\end{pmatrix}.
$$

(8.20)

For simplicity we assume that the observation mode matches perfectly the laser beam:

$$
\hat{E}^l(r) = \sqrt{\frac{2}{\bar{c}\varepsilon}} X(r) \hat{e} e^{ikz}
$$

(8.21)

$$
\hat{E}^o(r) = \sqrt{\frac{2}{\bar{c}\varepsilon}} X(r) \hat{e} e^{ik(z-d)}
$$

Note that we choose a different phase reference plane for $\hat{E}^l$ and $\hat{E}^o$. (In this way we include propagation phasor of the undisturbed laser beam in $T^t$.) The laser field $|E^l|$ is perfectly coupled into the receiver beam, but the receiver also picks up the forwardly scattered field $|E^s|$. In the transfer matrix, the superposition is expressed as

$$
T^t = I_2 e^{ikd} + T^s_{lo}.
$$

(8.22)

Here $I_2$ is a two-dimensional identity matrix that represents the perfect coupling of the undisturbed laser beam. The phasor $\exp(ikd)$ accounts for the propagation over
the distance \(d\). The matrix \(T_{lo}^S\) is the usual scattering transition matrix, whose matrix elements are given by Eq. 8.4. However, now we are interested in the special case \(\theta = 0\), i.e., \(q = 0\). This is a simplification, first of all because at \(\theta = 0\) the polarization matrix \(\Pi_d\) simplifies to \(I_2\), which means that the two diagonal elements of \(T_{lo}^S\) are equal: \(T_{xx}^S = T_{yy}^S = T^S\). The second simplification is \(q = 0\) in the integral in Eq. 8.4. Exploiting these simplifications, one finds that each row of Eq. 8.20 can be written as

\[
E^t(t) = [1 + i A(t)] e^{ik d} E^l,
\]

where \(A(t)\) is the forward scattering amplitude:

\[
A(t) = \frac{k \alpha}{2 \varepsilon_0} \int_{z=0}^d \int_x \int_y |X(x,y)|^2 c_m(r,t) dx dy dz.
\]

The molecular distribution \(c_m(r,t) = \bar{c} + \delta c(r,t)\) is a fluctuating quantity, and therefore also \(A(t)\) fluctuates. However, when looking in forward direction at \(q = 0\) the fluctuations are simply integrated over the IOR, and thereby more or less averaged out. Thus, in a good approximation we may replace \(c_m(r,t)\) with the time average \(\bar{c}\). Recalling that \(X(x,y)\) is quadratically normalized (Eq. 2.39), the pre-averaged forward scattering amplitude is readily evaluated and the pre-averaged version of Eq. 8.23 reads

\[
E^t = \left(1 + i \frac{k d}{2 \varepsilon_0} \bar{c}\right) e^{ik \Delta z} E^l
\]

Notice here a serious problem: Eq. 8.25 is apt to violate energy conservation!! When increasing the slab thickness \(d\), the forward scattering term will eventually dominate and grow towards infinity. This is because we neglected multiple scattering, i.e. the contribution of the secondary dipole fields to the field that drives the molecular oscillators. In forward direction the secondary waves are in phase, and thus amplified by constructive interference. We may neglect the secondary waves propagating backwards and sideward, but we are not allowed to neglect them in forward direction. In order to mend the problem, we sub-divide the slab into \(N\) thin layers of thickness \(\Delta z = d/N\). Then, starting from the first layer, we include the forward scattered field from one layer in the driving field of the next layer. Mathematically, this is expressed as

\[
E^t(t,d) = \Pi_{j=1}^N \left[1 + i A_j(t, \Delta z)\right] e^{ik d} E^l
\]

If we neglect fluctuations and assume that the medium is statistically homogeneous, then all the terms in the product are equal, and we can as well write

\[
E^t(d) = [1 + i A(\Delta z)]^N e^{ik d} E^l
\]

Here we insert \(A(\Delta z) = k \Delta z \bar{\chi}/2\), where \(\bar{\chi} = \bar{c} \alpha / \varepsilon_0\) is the mean susceptibility of the medium. Since \(\Delta z = d/N\), we remember that \(\lim_{N \to \infty} (1 + 1/N)^N = e\), which yields:

\[
E^t(d) = e^{ik \Delta z / 2} e^{ik d} E^l = e^{ik d [1 + \bar{\chi}/2]} E^l
\]

Since we expect that \(\bar{\chi} \ll 1\), we can as well write \(1 + \bar{\chi}/2 = \sqrt{1 + \bar{\chi}} = \sqrt{e} = n\). Thus, finally we obtain

\[
E^t(d) = e^{inkd} E^l
\]

The argumentation was somewhat “handwaving”, but it worked: the dilute scattering cloud behaves as an effective medium with refractive index \(n > 1\). But where is the attenuation? It is not readily visible in Eq. 8.29, because so far we neglected the
fact that the molecular polarizability $\alpha$ is a complex quantity: The driven molecular oscillators keep losing energy by radiation and therefore there is a phase shift between the driving force $E^d \exp(-i\omega t)$ and the response, i.e. the oscillating dipole $d(t) = \alpha E^d(t)$. In other words, we neglected the imaginary part of the polarizability $\alpha$. The correct expression for the polarizability of a radiation damped oscillator is:

$$\frac{\alpha}{\epsilon_o} = i \left[ 1 - e^{i\delta} \right] \frac{3\pi}{k^3},$$

(8.30)

where $\delta$ is the so-called scattering phase (nothing to do with the “phase function” of radiative transfer), i.e. the phase shift of the radiation damped dipole oscillator with respect to the driving field. A simple algebra exercise yields the real and imaginary parts of the polarizability as

$$\Re \frac{\alpha}{\epsilon_o} = \frac{3\pi}{k^3} \sin(\delta) \quad \Im \frac{\alpha}{\epsilon_o} = \frac{3\pi}{k^3} [1 - \cos(\delta)]$$

(8.31)

Upon inserting the complex susceptibility $\chi = \bar{c} \alpha/\epsilon_o = \bar{\chi}_{re} + i \bar{\chi}_{im}$ into Eq 8.29, we obtain

$$E^f(d) = e^{i n_{re} kd - kd \bar{\chi}_{im}/2} E^l,$$

(8.32)

where $n_{re} = \sqrt{1 + \bar{\chi}_{re}}$. Thus, by squaring the amplitude, we recover the Beer’s law:

$$P_l(d) = e^{-k \bar{\chi}_{im} d} P_l$$

(8.33)

The imaginary part of the susceptibility, that enters through the imaginary part of the forward scattering amplitude $A$, is responsible for extinction. This is the essence of the so called optical theorem [4]. Usually one does not care about deep physical reasons of extinction and the Beer’s law is formulated in terms of a measurable quantity, the so-called scattering coefficient $\mu_s$:

$$P_l(d) = e^{-\mu_s d} P_l = e^{-d/\ell} P_l \quad \text{where} \quad \mu_s = k \bar{\chi}_{im}$$

(8.34)

The inverse $\ell = 1/\mu_s$ is the scattering mean free path, see Fig. 8.3.

Radiation is not the only way to lose energy of the molecular oscillator. Most often scattering is accompanied by absorption. Absorption can be included by an additional damping term in the oscillator equation, but a rigorous treatment requires quantum physics. We content ourselves with the corresponding modification of the Beer’s law, i.e., we introduce the extinction coefficient $\mu_e$ (sometimes called “total attenuation coefficient”, $\mu_t$), that is the sum of $\mu_s$ and of the absorption coefficient $\mu_a$:

$$\mu_e = \mu_s + \mu_a$$

(8.35)

We shall also use the extinction length $\ell_e = 1/\mu_e$. Since absorption is quite easily included in Monte Carlo simulations, we forget about $\mu_a$ until the proper place.

### 8.2.1 Scattering coefficient and scattering cross-section

A particularly enlightening interpretation of the scattering coefficient is obtained if one realizes that

$$\Im \frac{\alpha}{\epsilon_o} = \frac{k^3}{6\pi} \frac{|\alpha|^2}{\epsilon_o^2} = \frac{1}{k} \sigma_d,$$

(8.36)

where $\sigma_d$ is the total scattering cross-section of an isotropic dipole oscillator. Then

$$\mu_s = \sigma_d \bar{c}.$$

(8.37)
The scattering coefficient turns out to be the \textit{density of scattering cross-section}. This interpretation, as well as Eq. 8.34, hold for an arbitrary scattering system. Note that the scattering coefficient from Eq. 8.37 is a purely incoherent superposition of the contributions of individual particles. There is no trace of inter-particle interference. This is because interference effects other than constructive forward interference concern fluctuations and those we neglected when deriving Eq. 8.37. The rigorous inclusion of fluctuation is a demanding exercise and we content ourselves with a self-consistent approximation based on energy conservation.

According to Beer’s law, the power removed from the beam upon passing a volume element $\delta V$ is $\delta P = \mu_s I \delta V$. Energy conservation requires that power removed equals the power radiated from this element by scattering. Energy conservation can be expressed as

$$\mu_s I \delta V = \int i_\Omega d\Omega \delta V$$  \hspace{1cm} (8.38)

where $i_\Omega$ is the \textit{density} of the radiant intensity [W sr$^{-1}$ m$^{-3}$], i.e., $i_\Omega \delta V$ is the radiant intensity emanating from a volume element $\delta V$ irradiated by $I$. The \textit{differential} radiometric version of Eq. 8.19 is:

$$i_\Omega(\theta, \phi; r) = \{\sigma_d c p_d(\theta, \phi) S(\theta, \phi)\} I(r)$$  \hspace{1cm} (8.39)

As usually, $I(r) = R Y(r)/\pi a_l^2$ is the local irradiance [W/m$^{-2}$] in the laser beam at the position $r$ in the sample. The quantity in curly brackets can be identified as the \textit{density of differential scattering cross-section}, abbreviated with $\{\partial \sigma\}$. Upon introducing $i_\Omega$ from Eq. 8.39 into Eq. 8.38, we find

$$\mu_s = \sigma_d c \int p_d(\theta, \phi) S(\theta, \phi) \sin(\theta) d\theta d\phi \quad p_R(\theta, \phi) = \frac{p_d(\theta, \phi) S(\theta)}{\int p_d(\theta, \phi) S(\theta, \phi) \sin(\theta) d\theta d\phi}$$  \hspace{1cm} (8.40)

Here $S(\theta)$ is the structure factor, $p_d(\theta, \phi)$ is the phase function of an isolated dipole (Eq. 7.44) and $p_R$ is the phase function of interacting Rayleigh scatterers. These results are prototype for scattering problems in the Born approximation, i.e., in the absence of multiple scattering other than in forward direction.

\section*{8.3 From dilute gas to continuum and back to Born approximation}

At this point we can make an attempt to be more specific about the meaning of “absence of multiple scattering” and about the validity of the Born approximation: It depends on the size $d_{IOR}$ of the region of interest: one is in the single scattering regime if Beer’s law, Eq. 8.33, applies for light beams traversing a path length $d_{IOR}$ in arbitrary direction through the sample (for illustration see Fig. 8.3). The criterion for the applicability of Born approximation is more strict. In order to use in Eq. 8.1 the unperturbed laser beam, we must require that $d_{IOR} \ll \ell$. Born approximation, in the strict sense as the absence of elementary multiple scattering, is not fulfilled in condensed matter. Even in such dilute gas as air at normal pressure the molecular concentration is as high as $10^8$ molecules per $\mu$m$^3$. The molecules are so close to each other that the electromagnetic interaction is dominated by the near field contribution $E_{near} \sim d/(\epsilon_o r^3)$ (recall Eq. 2.42), where $r \approx \bar{c}^{-1/3}$ is the mean intermolecular distance, and $d = \alpha E_o$ is the induced dipole moment. The elementary Born approximation requires that $E_{near} \ll E_l$, which means that $\alpha \bar{c}/\epsilon_o = \bar{\chi} \ll 1$. This condition is well
fulfilled in air where $\bar{\chi} = 0.0006$, but certainly not in water with $\bar{\chi} = 0.77$. In water one expects a significant contribution from multiple scattering, that is mediated by the near fields, instead of propagating photons. Fortunately, and somewhat paradoxically, this near field multiple scattering makes it possible to return to the Born approximation on a higher level: upon increasing the density, we end up with an electromagnetic quasi-continuum, whose susceptibility $\chi$ does no longer reflect the polarizabilities of the individual molecules, but is a collective property of the molecules together with their neighbors. A rigorous theoretical description of the transition to electromagnetic continuum is a demanding task (see, e.g., [39]), but the practical consequence is favorable: the results as developed so far remain applicable as long as the fluctuations of the susceptibility are sufficiently small so that multiple scattering from the fluctuations can be neglected. The only modification we must make is to replace the vacuum permittivity $\varepsilon_o$ with the effective medium permittivity $\varepsilon_m = \varepsilon_o(1 + \bar{\chi})$ and to multiply the vacuum wave number $k$ with the corresponding refractive index $n$ of the medium. The “unperturbed laser beam” is then to be understood as propagating in this homogeneous effective medium. A practical criterion for the applicability of the higher level of the first order approximation is the Beer’s law: the Born approximation becomes again applicable, if Beer’s law applies over a length scale that is much larger than the region of interest. What is “region of interest” depends on the context. The region of interest can be a sample cell, the IOR of a scattering experiment, or the structural correlation volume of size $\lambda_C$. In any case one must require $1/\ell = \mu_s \ll k$, which, according to Eq. 8.34, is equivalent to $\chi_{im} \ll 1$. This condition is well fulfilled with clean water whose scattering coefficient $\mu_s$ is as small as 0.0004m$^{-1}$ at 800nm [40]. In fact, susceptibility fluctuations in water are sufficiently small to neglect them in present context, and regard water as a homogeneous medium. In this medium we suspend sophisticated species of highly structured molecules: proteins, polysaccharides, hydrophobic fats, amphiphilic phospholipids. All these constituents will interact in manifold ways and self-organize into intriguing structures. We are about to create a biological tissue...

![Figure 8.3: Experiment demonstrating scattering regimes in a particle suspension with decreasing scattering mean free path $\ell = 1/\sigma c$.](image)

0: Reference, 1: Single scattering, Beer’s law, 2: Few scattering, still Beer’s law, 3: Multiple scattering

60
Chapter 9

Scattering from biological tissues

9.1 Scattering from particulate systems

Microstructures in biological tissue range from organelles 0.2-0.5 µm or smaller, to mitochondria 1-4 µm in length and 0.3-0.7 µm in diameter, to nuclei 3-10 µm in diameter, to mammalian cells 10-30 µm in diameter. The refractive index variation is about 0.04-0.10 for biological tissue with a background refractive index of $n_0 = 1.35$. [36]

In the present chapter we take the “tissue creation” seriously. We begin with a primordial soup containing the tissue constituents at low concentrations. The constituents of biological matter exhibit a strong tendency to aggregation that inevitably occurs upon increasing their concentration in water above a certain level. Phospholipids, as any surfactants, form micelles or vesicles, fats aggregate into droplets coated by a surfactant layer, amino-acids polymerize into peptides and protein macromolecules. These aggregates may be classified into four categories: R as Rayleigh, D as Debye, G as Gans and M for Mie.

Rayleigh particles: Aggregates of type R are small objects whose sizes $a$ satisfy the inequality $ka \ll 1$. This small size has two consequences: i) the internal structure of the aggregates can’t be resolved by light scattering, whose shortest probing length $1/q$ is $1/2k$; ii) they respond to electric field oscillations as simple dipoles with polarizability $\alpha$. This so-called Rayleigh scattering was the subject of chapters 7.6 and 8. The only slight difference is that now we regard $\alpha$ as the excess polarizability, resulting from the mismatch of refractive index of the particle and the surrounding medium. For example, the excess polarizability of spherical fat droplets in water is well approximated by the Clausius-Massoti-Lorentz-Lorenz-formula

$$\frac{\alpha}{\epsilon_w} = 4\pi \frac{m^2 - 1}{m^2 + 2} a^3,$$

where $\epsilon_w = 1.33^2 \times \epsilon_o$ is the dielectric constant of water and $m = n/n_w$ is the relative refractive index of the particle.

Debye particles: Aggregates of type D are large but fluffy objects, such as coils formed by linear macromolecular chains in good solvents. Since the coils are fluffy, we can neglect internal multiple scattering. Only the polarizabilities of the polymer units are slightly modified with respect to isolated monomers because of the near field coupling with their neighbors in the chain. Thus the scattering properties can be calculated using the Born approximation. However, the characteristic size $a$ of the coils is large, so that $ka \geq 1$. Therefore one must take into account interferences of the dipole waves radiated by the individual segments of a single coil. These interferences are there, even if the polymer suspension is highly diluted.
Gans particles: Aggregates of type G are large compact objects whose refractive index differs only slightly from the surrounding medium. In other words: \( ka \geq 1 \) but \( m \ll 1 \). The scattering amplitude from such aggregates can be modeled as the coherent superposition of the contributions from internal dipoles driven by the external field \( E_l^\prime \). This is quite similar to the Debye case, the only difference is of technical nature: in the Debye case one would use summation of contributions from discrete segments with polarizabilities \( \alpha_s \), whereas with Gans particles one regards the interior of a particle as a continuum with an excess susceptibility \( \chi_p \).

RDG regime: The common feature of the RDG particles is that we may disregard intra-particle multiple scattering: the field acting on their molecular constituents is the external field \( E_l^\prime \), e.g. the field of the illuminating laser, as it propagates through the homogeneous medium. In optics this is known as the Rayleigh-Debye-Gans-Born-approximation (RDGB).

Mie particles: Outside the RDG regime the matters become quite complicated. Relatively simple expressions for the particle scattering amplitudes of simple homogeneous objects such as spheres and ellipsoids have been provided by Mie, Gans and followers, but even those need intense numerical calculations. More complicated shapes require \textit{ab initio} numerics, for example the so-called T-matrix method. Fortunately, a biological tissue is not a dilute suspension of Mie particles, this does not even work for blood [41]. Therefore we can leave the subject “Mie and beyond” to standard texts [38, 24, 42, 43]. We shall venture into the Mie regime only briefly, in order to obtain a quantitative estimate of the applicability of the RDGB-approximation.

9.1.1 RDGB-scattering

Consider a single RDG particle, such as a protein macromolecule. For the sake of generality we shall regard the macromolecule as a small fluctuating lump of a dielectric quasi-continuum, which scatters light because of its excess susceptibility \( \chi_m \) with respect to the water background \( \chi_w \). The molecule is represented by the excess susceptibility distribution \( \chi_m[r-r_m; t] \), where \( r_m \) denotes the position of the molecule, perhaps the centroid of \( \chi_m(r) \). For simplicity we keep the position constant and assume that the size of the macromolecule, i.e. the size of the region where \( \chi_m > 0 \), is small with respect to the width \( a_l \) and \( a_o \) of the illumination and observation profiles. However, we allow the orientation, the form and the structure of the particle to fluctuate, in order to cover the case of flexible chain-like macromolecules.

In chapter 7.1 we worked out the formal theory for a general particle illuminated and observed by a quasi-plane wave and now we can earn the fruits. The only difference is that now we employ Born approximation, which means that the non-local susceptibility tensor \( \epsilon_w \mathcal{X}(r, r') \) is replaced by the local excess susceptibility \( \epsilon_w \chi(r) \delta(r-r') \). The subscript \( w \) is a reminder of the fact that now we work in water with susceptibility \( \tilde{\chi}_w \), which also enters that wavelength \( \lambda_w \) and all derived quantities. To simplify the matters even further, we assume isotropy of the response so that \( \epsilon_w \chi(r) \delta(r-r') = \epsilon_w \chi_m(r) \delta(r-r') I \), where \( I \) is the unit matrix. This expression we insert into Eq. 7.15 and calculate the polarizability tensor of the molecule:

\[
A_m(q, t) = I \int_V e^{i q \cdot u} \epsilon_w \chi_m(u, t) d^3u,
\]

where \( u = r - r_m \). The next step consists in two definitions:

\[
\alpha_m = \int \epsilon_w \chi_m(u, t) d^3u \quad \mathcal{F}(q, t) = \frac{1}{\alpha_m} \int e^{i q \cdot u} \epsilon_w \chi_m(u, t) d^3u
\]
Here $\alpha_m$ is the excess polarizability of the macromolecule. Note that $\alpha_m$ is the sum of the polarizabilities of the molecular constituents of the macromolecule and therefore approximately constant in time. The quantity $F(q, t)$ is the so-called form amplitude. The polarization transfer matrix of the RGD-particle is that of an elementary dipole. Therefore the S-matrix is essentially a copy of Eq. 7.42, the only new feature is the form amplitude $F(q, t)$:

$$ S^{RDG}(t) = \frac{i k_c^3 \alpha_m}{4 \pi \epsilon_w} F(q_w, t) \begin{pmatrix} 1 & 0 \\ 0 & \cos(\theta) \end{pmatrix}. \tag{9.4} $$

The form amplitude of a general RGD-particle is a fluctuating quantity. The measured signal depends on its average $F(q) = \langle |F(q, t)|^2 \rangle$, which is called form factor, akin to the structure factor $S(q)$ from Sect. 8.1:

$$ F(q) = \langle |F(q, t)|^2 \rangle = \left| \frac{\epsilon_w}{\alpha_m} \right|^2 \int \int e^{i q_u \cdot (u-u')} \gamma_m(u, u') d^3 u d^3 u'. \tag{9.5} $$

where $\gamma_m(u, u') = \langle \chi_m(u, t) \chi_m(u', t) \rangle$ is the susceptibility correlation function that is akin to the correlation function $g_6(\rho)$ of density fluctuations from Eq. 8.11. In terms of the form factor, the normalized time-averaged signal from a single particle is:

$$ \langle P_{rgd}^1 \rangle_{\theta>0} = \frac{1}{\pi a_l^2} \left[ \frac{k_w^4}{6 \pi} \left| \frac{\alpha_m}{\epsilon_w} \right|^2 F(q_w) p_d(\theta, \phi) \right]\tag{9.6} $$

Here $p_d(\theta, \phi)$ is the phase function of an isolated dipole, Eqs. 7.44 and 7.45. The bracketed quantity is the differential scattering cross-section $\partial \sigma_{rgd}$ of the RGD-particle. Correspondingly, the total scattering cross-section is the integral

$$ \sigma_{rgd} = \frac{k_w^4}{6 \pi} \left| \frac{\alpha_m}{\epsilon_w} \right|^2 \int p_d(\theta, \phi) F(\theta, \phi) \sin(\theta) d\theta d\phi \tag{9.7} $$

Finally, the phase function $p_R$ of RGD scatters:

$$ p_{rgd}(\theta, \phi) = \frac{p_d(\theta, \phi) F(\theta, \phi)}{\int p_d(\theta, \phi) F(\theta, \phi) \sin(\theta) d\theta d\phi} \tag{9.8} $$

Thus, Eq. 9.6 can be abbreviated as

$$ \langle P_{rgd}^1 \rangle_{\theta>0} = \frac{\Omega_w \sigma_{rgd}}{\pi a_l^2} p_{rgd}(\theta, \phi). \tag{9.9} $$

Obviously, it is quite difficult to keep in the IOR a single particle. Usually one works with a suspension of a certain number density $\bar{c}$. If the suspension is sufficiently dilute so that interactions between the particles can be neglected, then

$$ \langle P_{rgd} \rangle_{\theta>0} = \Omega_w \left\{ \frac{k_w^4}{6 \pi} \left| \frac{\alpha_m}{\epsilon_w} \right|^2 F(q_w) p_d(\theta, \phi) \right\} V_{IOR} = \Omega_w \left\{ \sigma_{rgd} p_{rgd}(\theta, \phi) \bar{c} \right\} V_{IOR}. \tag{9.10} $$

Notice in the curly brackets the density of differential scattering cross-section $\{ \partial \sigma_{rgd} \}$ of a dilute suspension of RGD particles. With increasing $\bar{c}$ interactions set in and so we would have to return to Sec. 8 and re-do the work with RGD particles in place of the elementary dipoles. However, since a dilute suspension is anyway a bad model of a biological tissue, we content ourselves with Eq. 9.10.
9.1.2 Interpreting form factors

For future use we compile here the form factors of basic particle shapes as they can be found in many texts on small angle scattering. On the right hand sides of the following equations we indicated the limiting behavior for large $q$, when probing length scales that are much shorter than the size of the particle. To make this limiting behavior visible, the form factors in Fig. 9.1 are plotted in log-log-representation. All formulas assume that the scatterers are isotropic in the statistical sense, for example, because of free random rotation. Then $F(q)$ depends only on $q = |q| = k \sqrt{2 - 2 \cos(\theta)}$.

The upper curve in Fig. 9.1 represents homogeneous rods of length $L$:

$$F_{\text{rod}}(q) = \frac{2}{qL} \left[ \text{Si}(qL) - \frac{1}{qL} \cos(qL) \right] \rightarrow \frac{1}{q^2}$$  (9.11)

The $q^{-1}$ law at large $q$ reflects the fact that a thin rod is a 1-dimensional object.

Thin homogeneous disc-like sheet of radius $R$:

$$F_{\text{disc}}(q) = \frac{2}{q^2 R^2} \left[ 1 - \frac{J_1(2qR)}{2qR} \right] \rightarrow \frac{1}{q^2}$$  (9.12)

Again, $q^{-2}$ law at large $q$ reflects the dimensionality of the homogeneous structure.

Thin spherical shell of radius $R$. The shell form-factor may be of a special interest because spherical vesicles are a good models of cell membranes:

$$F_{\text{disc}}(q) = \frac{\sin^2(qR)}{q^2 R^2} \quad \text{(smoothed)} \rightarrow \frac{1}{q^2}$$  (9.13)

Neglecting the oscillation, which are anyway smoothed out in a polydisperse system, one again sees the $q^{-2}$ behavior. This indicates the 2D sheet of which the shell membrane is made. Another frequently encountered object is a sphere, for example the nucleus of a cell. Homogeneous sphere of radius $R$:

$$F_{\text{sph}}(q) = \left\{ \frac{3[\sin(qR) - qR \cos(qR)]}{(qR)^3} \right\}^2 \rightarrow \frac{1}{q^4}$$  (9.14)

The simple dimensionality rule breaks down at dimension 3: there is no way to look into a homogeneous 3D structure (at large $q$) because of the lack of scattering. Only
surfaces of such 3D object are visible. \( q^{-4} \) is characteristic for randomly oriented smooth interfaces. According to Porod and Debye, any two-phase system with smooth interfaces between the phases will scatter so that \( F(q) \sim q^{-4} \) at large \( q \).

We also consider two objects of fractal character. Fractals are structures that exhibit a certain degree of self-similarity [44]. When investigating the object on shorter and shorter length scales (i.e. at larger and larger \( q \)), and adjusting the magnification accordingly, one keeps finding the same picture, at least in the statistical sense. In practice, however, there are two cut-off lengths, \( L \) and \( l \), above and under which the fractal picture breaks down: large \( L \) is the overall size of the object and small \( l \) is the size of its constituent monomers. A classical fractal is a random polymer coil with a characteristic size \( R_g \):

\[
F_{\text{coil}}(q) = \frac{2[\exp(-q^2R_g^2) + qR_g - 1]}{q^2R_g^2} \quad \rightarrow \quad \frac{1}{q^2} \quad (9.15)
\]

When magnifying the polymer coil one always sees the same randomly entangled loops. The coil is a “mass fractal”. It is a 3D object, but it does not fill the 3D space. The fractal dimension of the coil is 2, and therefore \( q^{-2} \). But the same large \( q \) behavior exhibits the 2D disk or shell. It is not easy to identify a true fractal [45]. Therefore we shall not dwell on the fractal point. For future use we record only one more form factor, namely one that belongs to an UFO (Unidentified Fractal Object) that is also known under the name “HG”:

\[
F_{\text{HG}}(q) = \frac{1}{(1 + q^2L^2)^{3/2}} \quad \rightarrow \quad \frac{1}{q^3} \quad (9.16)
\]

An advanced reader interested in fractals will have noticed that this form factor is exactly on the border line between a mass fractal and a surface fractal.

Common to all six form factors is their limiting behavior at \( q \rightarrow 0 \), in the so-called Guinier regime at nearly forward scattering angles \( \theta \). Expanding Eq. 9.5 in a Taylor series in \( q \), one finds that

\[
F(q) = 1 - \frac{q^2R_g^2}{3} + O(q^4R_g^4) \quad \text{where} \quad R_g = \frac{\int \int \rho^2 \gamma_m(u,u + \rho) d^3ud^3\rho}{\int \int \gamma_m(u,u + \rho) d^3ud^3\rho} \quad (9.17)
\]

Here \( \gamma_m(u,u') = \langle \chi_m(u,t)\chi_m^*(u',t) \rangle \) is the susceptibility correlation function. \( R_g \) is the so-called radius of gyration, an universal measure of the size of the scattering object. All form factors in Fig. 9.1 correspond to the same \( R_g \).

### 9.1.3 On the applicability of the RDGB-approximation

A single isolated Mie sphere may not be a good model for light scattering in tissue, but since Mie algorithms are readily available\(^1\), we can use it as a simple model to estimate the applicability of the Born approximation for modeling the optical properties of tissues. We consider a rather large sphere with a diameter \( 2a \approx 3.3\mu m \), which gives the size parameter \( ka = 17.3 \) when using a 800nm laser. Refractive index of the sphere is 1.42 and the surrounding medium is water. The sphere may represent the nucleus of a mammalian cell. In Fig. 9.2 we plot the Mie intensity \( I(q) \sim (|S_1|^2 + |S_2|^2) \) together with the RDG-form factor from Eq. 9.14. Such a large sphere scatters predominantly forward, with an asymmetry parameter \( g = 0.98 \) (recall Eq. 7.32). In the linear plot the difference between Mie and RGD is hardly discernable and even at large angles,

\(^1\)We use an IDL version of Bohren-Huffman code [24] by B.T. Drain and P.J. Plateau.
that are better visible in the log-log plot, the deviations appear to be marginal. Less good, but still reasonable, is the agreement of the total differential cross-section. The exact Mie result is \( \sigma_M = 2.12 \times \pi a^2 \), whereas the RDG approximation, Eqs. 9.1 and 9.7, yields \( \sigma_{RDG} = 2.64 \times \pi a^2 \). Since we are interested in polarization, we also plotted in Fig. 9.3 the ratio \( |S2/S1|^2 \). In Born approximation one expects \( |S2/S1|^2 = \cos^2(\theta) \). The corresponding Mie curve is considerably wilder, indicating a resonant excitation of higher order multipoles. Nevertheless, even in the Mie range, the dipolar character of the polarization remains to some extent preserved. Except for the resonance peaks, that are easily smoothed out by size averaging, there is only a slight shift of the minimum towards large angles. Because of the strong forward anisotropy of the scattering, the large angles play anyway a minor role. Thus, we conclude: when modeling the optical properties of tissues, it is hardly worth to bother with Mie spheres or more complicated Mie objects. RDGB approximation should be sufficient in most cases.

![Figure 9.3: Ratio \(|S2/S1|^2\) of a Mie sphere (solid line) and a RDG sphere (dotted line).](image)

**9.2 Scattering from a Continuum**

While propagating through a tissue, a light wave can encounter structures with dimensions ranging from the width of a protein macromolecule (50-100 nm), to the diameter of the body of a cell (2-50 \( \mu m \)), to the length of a blood vessel (.1 mm). Moreover, the constituents of tissues are tightly compacted, with their surfaces pressed together to form contiguous units. Theories based on incoherent scattering among discrete particles cannot be applied at the microscopic level in tissue without violating their underlying assumptions. [46]
Equipped with the knowledge about scattering from dilute particulate systems we return to the tissue creation laboratory. We increase the density of bio-matter in water to realistic volume fractions of about 10-20%. At such a high density, the bio-particles, such as membrane sheets, collagen rods and coils, protein spheres and coils, find themselves well within the range of their interaction forces. The picture of independently scattering particles ceases being meaningful. Thus, we revert to the continuum picture of a medium with fluctuating susceptibility $\chi(r, t)$. We have already done most of the work and so we only need to recapitulate the results: i) scattering at finite angles is due to fluctuations $\delta \chi(r, t)$. ii) forward scattering is included in the scattering coefficient $\mu_s$, and in the mean refractive index $n_m$ of the medium. This $n_m$ corresponds to the real part of the mean susceptibility $\bar{\chi}$. Equipped with this knowledge, we return to Eq. 8.5 but express the matrix element $T^{1}_{lo}$ in terms of the susceptibility fluctuations instead of molecular concentration:

$$T_{dl}(t) = e^{ik_m L_{lf}} e^{-\mu_s L_{lf} / 2} \frac{k_m}{2} \int X_o(r) X_i(r) \delta \chi(r, t) e^{i \mathbf{q_m} \cdot \mathbf{r} \hat{d} \mathbf{r}} (\hat{e}_t \cdot \hat{e}_o)$$  \hspace{2cm} (9.18)

The subscript $m$ at $k_m$ and $q_m$ is a reminder of the fact that now we work in an effective medium with susceptibility $\bar{\chi}_m$, that also enters the wave number and wavelength $\lambda_m$. When calculating the time averaged signal that is proportional to $\langle |T_{dl}(t)|^2 \rangle$, one naturally encounters the correlation function of susceptibility fluctuations:

$$\gamma_\delta(\rho) = \langle \delta \chi(r, t) \delta \chi^*(r + \rho, t) \rangle$$  \hspace{2cm} (9.19)

We assume statistical homogeneity within the IOR, i.e., $\gamma_\delta(r, r + \rho) = \gamma_\delta(\rho)$ depends only on the displacement $\rho = r' - r$. In the continuum model, $\gamma_\delta(\rho)$ replaces the correlation function $g_\delta(\rho)$ of density fluctuations of the interacting particle model in Sec. 8; recall Eq. 8.11. Recall also that $g_\delta(\rho)$ diverges for $\rho = 0$, there is the self-correlation term $\delta \delta(\rho)$. A similar behavior can be expected for $\gamma_\delta(\rho)$. Thus the fluctuation variance $\gamma_\delta(0) = \langle |\delta \chi|^2 \rangle$ is not a good measure of the strength of scattering, despite erroneous statements in some texts. As it turns out, a better measure of the strength of scattering is the integral of the correlation function

$$V_\chi = \lambda_\chi^2 = \int \gamma_\delta(\rho) d^3 \rho$$  \hspace{2cm} (9.20)

Note that susceptibility is a dimensionless quantity and therefore $V_\chi$ can be interpreted as the susceptibility correlation volume. Likewise, $\lambda_\chi$ is the susceptibility correlation length, i.e. the range of $\rho$ where $\gamma_\delta(\rho)$ is non-zero. Having reflected on the properties of the susceptibility correlation function, we calculate the normalized time averaged signal $\langle P_{os} \rangle$. Recall that $\langle P_{os} \rangle$ is the probability that a photon traveling in a laser beam from a plane at $r_i$ is scattered from the IOR centered at $r_f$ into the observation mode. Expressed in terms of $\gamma_\delta(\rho)$ the normalized signal reads

$$\langle P_{os} \rangle_{\theta > 0} = e^{-\mu_s L_{lf} / \Omega_m} \frac{k_m^4}{\pi a_i^2} \frac{1}{6 \pi} \int Y_{d}(r) \gamma_\delta(r', r) e^{i \mathbf{q_m} \cdot (r - r')} Y_{d}(r') d^3 r d^3 r' p_d(\theta, \phi)$$  \hspace{2cm} (9.21)

where $p_d(\theta, \phi)$ is the dipole phase function, which appears here, because the Born approximation preserves the dipole character of polarization transfer. If $\lambda_\chi$ is much shorter than the size of the IOR then we factor the integration and write

$$\langle P_{os} \rangle_{\theta > 0} = e^{-\mu_s L_{lf} / \Omega_m} \frac{k_m^4}{\pi a_i^2} \frac{1}{6 \pi} S_\chi(\mathbf{q}) p_d(\theta, \phi) V_{IOR}.$$  \hspace{2cm} (9.22)
where \( V_{IOR} = \int Y_o(\mathbf{r}) Y_l(\mathbf{r}) d^3r \) is the effective volume of the IOR and \( S_\chi(\mathbf{q}) \) is the spatial power spectrum of the susceptibility fluctuations, i.e., Fourier transform of the correlation function \( \gamma_\delta(\mathbf{r}) \):

\[
S_\chi(\mathbf{q}) = \int \gamma_\delta(\mathbf{r}) e^{i \mathbf{q} \cdot \mathbf{r}} d^3r
\]

Note that this \( S_\chi(\mathbf{q}) \) is analogous with the structure factor \( S(\mathbf{q}) \) from Eq. 8.18, but here the pre-factor \( 1/\bar{c} \) is missing. A convenient normalization that is appropriate for the present case is to use \( F(\mathbf{q}) = S_\chi(\mathbf{q})/V_\chi \); recall that \( F(0) = 1 \). So we write

\[
\langle P_{os} \rangle_{\theta>0} = e^{\mu_s L_{ij} \Omega_m \frac{1}{\pi d_l^2} \left\{ \frac{k^4}{6\pi} V_\chi F(\mathbf{q}) p_d(\theta, \phi) \right\}} V_{IOR}.
\]

9.2.1 Quasi-particles

When comparing Eq. 9.24 with the corresponding Eq. 9.10 for independently scattering RGD-particles, one finds only two differences. First difference is the attenuation factor \( \exp(-\mu_s L_{ij}) \), which we introduced to account for the effective medium’s complex susceptibility \( \bar{\chi} \); recall Sec. 8.2. The second difference is one replacement:

\[
V_\chi \leftrightarrow \frac{\alpha_m}{\epsilon_w} \bar{c}^2
\]

In the continuum model, the product of mean concentration and the square of particle polarizability is replaced with the susceptibility correlation volume \( V_\chi \). So it turns out that the RDGB theory of independent particle scattering is only a special case of light electrodynamics in Born approximation, slightly complicated by the “simplification” of having a dilute suspension of discernable isolated particles. Thus, there is no need to assume independent scattering. However, if one wishes, then one can imagine the IOR as being filled with independently scattering quasi-particles, whose scattering cross-section takes into account the correlations of the elementary scatterers with their neighbors. The size \( L_q \) of the quasi-particles is the correlation length \( \lambda_\chi \) of the susceptibility fluctuations. Such quasi-particles are stochastic ghosts, reflecting the power spectrum (from Latin spectrum, apparition, ghost) of susceptibility fluctuations. They exist only on average, much like laser speckles formed by scattering of coherent light.
on rough surfaces. Because a quasi-particle is only an average ghost, we place it in an average position in the center of IOR. If we accept the existence of quasi-particles, then one would perhaps like to define their effective concentration $\bar{C}$. We refrain from such attempt, since quasi-particles are only ghosts, and since their concentration anyway cancels in the final result. The measurable quantities are the scattering coefficient $\mu_s$ and the phase function $p(\theta, \phi)$.

### 9.3 Scattering matrices of fluctuating continuum

The scattering transition matrix $T_{lo}^I(t)$ of a particle suspension or of a sample of biological tissue is a randomly fluctuating quantity, not well suited for practical use in, say, polarized Monte Carlo simulation. However, thanks to the Born approximation $T_{lo}^I(t)$ is the product of a scalar complex amplitude $A(t)$ and a polarization matrix $\Pi^d$, which only concerns the polarization transfer:

$$T_{lo}^I(t) = e^{ikmL_{lo}}e^{-\mu_sL_{lo}/2}A(t)\Pi^d,$$

where

$$A(t) = \frac{i k_m}{2} \int X_o(r)X_l(r)\delta \chi(r, t) e^{i q_m \cdot r}d^3r \quad \Pi^d = \begin{pmatrix} 1 & 0 \\ 0 & \cos(\theta) \end{pmatrix}$$

(9.29)

Fluctuations of $A(t)$ concern only the angular distribution of the scattered light. The polarization state, that is operated on by the constant polarization matrix, does not fluctuate. In the Born approximation, the single scattering process does not depolarize, a pure polarization state is scattered into another pure polarization state. Therefore, in the practice of polarized Monte Carlo simulations we are allowed to employ the Jones formalism. Since we shall neglect interferences between different photon paths, it is sufficient to use the time or ensemble pre-averaged version of the transition matrix:

$$\bar{T}^I = i \sqrt{\frac{\mu_s V_{IOR}}{\pi a_l^2}} \Omega_o e^{ikmL_{lo}}e^{-\mu_sL_{lo}/2} \left\{ \sqrt{\frac{F(q)}{\Phi}} \sqrt{\frac{3}{8\pi}} \begin{pmatrix} 1 & 0 \\ 0 & \cos(\theta) \end{pmatrix} \right\},$$

(9.31)

where $F(q)$ is an experimentally determined or modeled structure factor and $\Phi$ is the already well known normalization constant

$$\Phi = \int F(q) p_d(\theta, \phi) \sin(\theta) d\theta d\phi$$

(9.32)

Recall that if the structure is statistically isotropic, then $F(q)$ depends only on $q = |q| = k\sqrt{2 - 2\cos(\theta)}$. This is a welcome simplification, but structural isotropy can’t be a priori expected in biological tissues. Most tissues exhibit both, structural anisotropy and anisotropy of dielectric response. We leave this difficult case for future work.

The quantity in large curly brackets can be recognized as the normalized S-matrix $\hat{S}$ that was defined in Eq. 7.34 as $\hat{S} = S/\sqrt{k^2 \sigma_t}$. Instead of the total scattering cross-section $\sigma_t$ we have now the density of the scattering cross-section $\mu_s$, but still, the phase function can be calculated as

$$p(\theta, \phi) = \frac{\langle E_l^l|\hat{S}^\dagger\hat{S}|E_l^l\rangle}{\langle E_l^l|E_l^l\rangle} = \frac{\langle E_l^l|R_2^{-1}\hat{S}^\dagger\hat{SR}_2|E_l^l\rangle}{\langle E_l^l|E_l^l\rangle}$$

(9.33)

If one wishes, one can insert into the curly bracket the normalized S-matrix of a Mie-sphere. Then one would interpret the scattering coefficient as $\mu_s = \sigma_t \bar{c}$, where $\sigma_t$ is the
scattering cross-section and \( \bar{c} \) the number density of the spheres. Everything is allowed in the world of “quasi-particles”.

We leave it to the reader to verify that \( T_1(t) \) from Eq. 9.29 and \( \bar{T}_I \) from Eq. 9.31 result in the same Mueller matrix \( M \). In this sense \( \bar{T}_I \equiv \sqrt{M(\phi, \theta)} \). Such behavior can be expected for all systems with isotropic dielectric response, in particular also for isotropic Mie spheres. Keep in mind, however, that the concept of the pre-averaged Jones matrix is not universally applicable. It would fail if the scattering system consists in an ensemble of anisotropic scatterers with randomly fluctuating orientation of the anisotropy axis \( \hat{a}(t) \). In this case, depolarization would occur and we would have to resort to Mueller formalism. This is not because “Jones formalism can’t treat unpolarized light”, but because “a Mueller matrix is equivalent to an ensemble of Jones matrices” [28]. Unfortunately, single scattering depolarization makes the treatment of anisotropy in multiple scattering rather difficult. To include anisotropy, we need to know the polarization vector \( e \), but Mueller formalism yields only averages of products of its complex components. This difficult case of anisotropic media with fluctuating anisotropy axis is discussed in Ref. [47].

### 9.4 Modeling the tissue structure factor

One way to approach the tissue optics would be to construct the susceptibility correlation function \( \gamma_\delta(r, r') \) and the corresponding structure factor \( F(q) \) from scratch, that is by \textit{ab initio} modeling of the tissue structure. Today, in the computer age, the task is not impossible. Attempts for such \textit{ab initio} modeling have been reported for example in [48, 49, 50]. However, one must realize that the modeling task is exactly the same as physically constructing the tissue itself. We let the nature to do the job and rely on empirical evidence. For a long time the work horse of tissue optics was the scalar radiative transfer, more recently in the Monte Carlo version. With scalar light one does not distinguish between the structure factor and the phase function. The most commonly employed model is still the Henyey-Greenstein phase function [51]:

\[
p(\hat{s}_1, \hat{s}_o) = p(\theta) = \frac{1 - g^2}{4\pi} \frac{1}{[1 + g^2 - 2g \cos(\theta)]^{3/2}} \quad \text{where} \quad g = \langle \cos(\theta) \rangle \quad (9.34)
\]

\( g \) is the asymmetry parameter. Since its introduction into tissue optics [52], the Henyey-Greenstein phase function was found to describe the angular dependence of single scattering from biological tissue reasonably well. Which is kind of surprising, because this phase function was originally introduced to model the “Diffuse radiation in the galaxy” [51]. In galactical as well as in tissue optics, the HG phase function is usually regarded as an empirical approximation for Mie scattering from independent particles with a distribution of sizes. This interpretation makes sense in the former branch of optics, but is questionable in the latter. Nevertheless, even in tissues HG remains a good and well accepted approximation.

A new era of tissue optics seems to have been initiated by Schmitt and Kummar in 1996 [46]. Using a phase contrast microscope they recorded the spatial variations of the optical thickness of tissue samples (geometrical thickness 5\( \mu \)m) and analyzed those variations in terms of two-dimensional spatial power spectra. They found the data to be well modeled with an expression used for modeling the light propagation through turbulent atmosphere:

\[
\tilde{\gamma}_2(q) \sim \frac{1}{[1 + q^2 L^2]^m} \quad (9.35)
\]
The exponents \( m \) found in four different types of tissues lay in a relatively narrow range 1.30-1.45 and cut-off length \( L_o \) was roughly 1\( \mu \)m. (Since in [46] spatial frequency is defined as \( \kappa = q/2\pi \), the \( L_o \) values in Table 1 in [46] must be divided by \( 2\pi \).) At large \( q \), so that \( qL_o > 1 \), the data are well described by a nice power law \( q^{-2m} \) (2\( m \approx 2.5 - 2.9 \)) that applies over more than three decades of the decay of \( \tilde{\gamma}_2(q) \). Similar results obtained with a similar technique have been published “for the first time” recently: in [53] one finds somewhat lower exponents and less nice power laws, but the general picture is the same. It is not quite obvious if such two-dimensional data, obtained with sample thickness of the same order of magnitude as \( L_o \), are indeed representative of the three dimensional distribution of susceptibility fluctuations. However, an independent confirmation of the power law behavior has been provided by Popp et al. [54]. These authors employed a light scattering microscope to measure directly the angular distribution of scattering intensity from thin tissue layers. Their data shown in Fig. 9.4 indicate power law exponents 2\( m \) in the range 2.3-2.7. (An exception is the skin: the exponent 2\( m \approx 3.6 \) approaches the exponent 4 of Porod’s law.)

Power law decays of power spectra, especially fractional power laws, are suggestive of fractals. The fractal idea triggered an intense burst of activities in optical modeling [55, 56, 49, 57, 36, 50, 58], but such modeling is beyond the scope of the present work. Fractals are a matter of taste. One can as well imagine an average structure (quasi particle) resulting from a polydisperse mixture of elementary shapes, rods, spheres, discs ... Consider for example the HG phase function. When translated into the present scattering language, the HG power spectrum reads

\[
\tilde{\gamma}(q) \sim \frac{1}{1 + q^2 L^2} \quad \text{where} \quad L^2 = \frac{1}{k^2 (1 - g)^2}, \quad (9.36)
\]

where \( L \) is the size of the quasi-particle, or the correlation length. Note that the asymptotic power law exponent of 3 is not too far from the experimental values 2.5-2.9. This may explain the long lasting success of the HG phase function. As already mentioned, the \( q^{-3} \) power spectrum can be interpreted as a border line fractal, either an extreme mass fractal or an extreme surface fractal. However, the same power spectrum can and has been interpreted as resulting from polydisperse Mie spheres with broad size distribution. (One can assume a power law [56].) Therefore, instead of promoting the fractal model, we content ourselves with a practical proposal. In order to acknowledge the new developments in the practice of radiation transfer in tissues, one should include
the “fractal” exponent as a parameter into the scattering law (i.e. in the structure factor). A simple but rather general and mathematically tractable model for the isotropic case is easily found, if one realizes that $q \tilde{\gamma}(q)$ and $\rho \gamma(\rho)$ are a pair of Fourier sinus-transforms. When skimming through the corresponding table in Gradshteyn Ryzhik [59], one immediately stumbles over the following formula:

$$\tilde{\gamma}(q) = \frac{L^3}{[1 + (Lq)^2]^{\nu+3/2}} \quad \text{where} \quad \nu > -1$$  \hspace{1cm} (9.37)

The appearance of the exponent $3/2$ is of course reminiscent of HG, which is recovered when $\nu = 0$. The extreme $\nu = -1$ would correspond to a 1D structure, recall the form factor of rods, Eq. 9.11. The corresponding correlation function reads:

$$\gamma(\rho) = \sqrt{\pi} \frac{(\rho/L)^\nu}{2^{\nu+1} \Gamma(\nu + 3/2)} K_\nu(\rho/L)$$  \hspace{1cm} (9.38)

Note that for $\nu \leq 0$ the correlation function diverges at $\rho = 0$. In other words, the variance of susceptibility fluctuation $\langle \delta \chi^2 \rangle$ appears to be infinite. This is because we did not include the short scale cut-off $l_m$, namely the finite size of the molecular constituents. This could be mended by convoluting Eq. 9.38 with a suitable molecular profile, for example with a Gaussian $\exp(-\rho^2/l_m^2)$, which would then multiply $\tilde{\gamma}(q)$ by the Gaussian $\exp(-q^2l_m^2)$. Not much would change because $l_m \ll L$; the molecular length scale is not observable in light scattering experiments. Recall that the scattering intensity is not proportional to $\gamma(0) \approx \langle \delta \chi^2 \rangle$ but to $\int \gamma(\rho) d^3\rho = \int \langle \delta \chi(0) \chi(r) \rangle d^3\rho$, which is finite. A more detailed discussion of the properties of the correlation function 9.38 can be found in a recent paper by Sheppard [58], with whom we share the practical proposal, without stressing the fractal issue.

Perhaps even more important than fractalization of the tissue micro-optics would be to develop a good model of anisotropy, which is present in most kinds of tissue. Two kinds of anisotropy can be expected to play a role. We already encountered the anisotropy of dielectric response, which is associated with birefringence. Dielectric anisotropy affects the dipole S-matrix, as already discussed in chapter 7.6. The second kind of anisotropy is the anisotropy of structure, leading to an isotropic correlation function and thus also to anisotropic light propagation. (Not to be confused with the $\theta$-asymmetry of the phase function.) A striking example is found in [54]. Usually, the structural and dielectric anisotropy are coupled. Both kinds of anisotropy can be incorporated in MC simulations, but this is a rather tiring exercise in geometry.

Remember, finally, that in order to apply the present version of the continuum scattering theory one must be able to find a region of interest that is much larger than the correlation length $\lambda_C \approx L$, but much smaller that the mean free scattering length $\ell = 1/\mu_s$. Thus, when determining the structure factor experimentally, one should always verify the existence of the single scattering limit. This is done by measuring with tissue slices with decreasing thickness $d$. As long as $d > \ell = 1/\mu_s$ one must expect the measurement to be polluted by multiple scattering so that the parameter values, say $L$ and $\nu$, are not reliable. Thus, one better takes a thinner slice and repeats the experiment, recording the new values. If the parameters did not change, then we are in the single scattering limit. If there is change, we repeat the procedure once again, and so on. Does the procedure converge? If not, then we are forced to revise the theory.
Chapter 10

Multiple scattering and radiative transfer.

In tissue optics, the “region of interest” is not a microscopic volume of the size of few micrometers, but a macroscopic sample, that is much larger than the scattering mean free path $\ell = 1/\mu_s$. We must take multiple scattering into account. The rigorous way would be to proceed to a second order approximation, which means adding to the laser field $E'$ the first time scattered field $E^{(1)}$, generated by interaction of $E'$ with the susceptibility fluctuations $\delta\chi$. This is the second step in the following hierarchy:

$$
\delta P^{(1)}(r, t) = \epsilon \delta \chi(r, t) E'(r) \quad \delta P^{(1)}(r, t) \rightarrow E^{(1)}(r, t) \quad (10.1)
$$

$$
\delta P^{(2)}(r, t) = \epsilon \delta \chi(r, t) [E'(r) + E^{(1)}(r, t)] \quad \delta P^{(2)}(r, t) \rightarrow E^{(2)}(r, t) \quad (10.2)
$$

The scheme is to be understood as follows: in the first order approximation we were concerned with the fluctuations of the dielectric polarization density $\delta P^{(1)}(r, t)$ as the result from interaction of the laser field $E'(r)$ (as it propagates in homogeneous media of susceptibility $\bar{\chi}$) with the susceptibility fluctuations $\delta\chi(r, t)$. This density of oscillating dipole moments corresponds to the current density distribution $j(r, t) = -i\omega \delta P^{(1)}$, which we then inserted into the Equation 5.6 in order to calculate the amplitude $E^0$ of the beam wave selected by the receiver from the scattered field. Now we postpone this step and instead use $\delta P^{(1)}(r, t)$ to calculate the first order scattered field $E^{(1)}(r, t)$ itself. This $E^{(1)}(r, t)$ is a superposition of spherical waves radiating from the dipoles that constitute the distribution $\delta P^{(1)}(r, t)$. Those dipole fields consist of near- and radiation-contributions, as given in Eqs. 2.42 and 2.43. Since we keep assuming that $k\ell = k/\mu_s \gg 1$, we can neglect the near-field, but the expression for $E^{(1)}(r, t)$ still looks difficult, even in the simplest case of isotropic dielectric response:

$$
E^{(1)}(r, t) = \frac{1}{4\pi} k^3 \int_V \frac{e^{i\mathbf{k} \cdot (\mathbf{r} - \mathbf{r}')}}{|\mathbf{r} - \mathbf{r}'|^2} \left[ E^{(l)}(r') - \frac{(\mathbf{r} - \mathbf{r}') \cdot (\mathbf{r} - \mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|^2} E^{(l)}(r') \right] \delta\chi(r') \, d^3r' \quad (10.3)
$$

This $E^{(1)}$ we would then insert into Eq. 10.2, the resulting $\delta P^{(2)}(r, t)$ in turn into Eq. 5.6, and then the task would be to re-do the work we have done so far. Thereby we would be facing rather ugly monsters, namely higher order fluctuation correlation functions, up to $\langle \delta\chi(r_1) \delta\chi(r_2) \delta\chi(r_3) \delta\chi(r_4) \rangle$. In view of such an outlook, we immediately interrupt the rigorous procedure. Since in our case $\ell = 1/\mu_s$ is much smaller than the size of the sample, we would anyway have to iterate the procedure great many times to achieve a realistic result. The iterative scheme may be straightforward, but the convergence is very slow and the procedure only works in few special cases [60]. The main problem
is the infinite range of the radiation interaction: all involved integrals must be evaluated over the macroscopically large volume of the sample. An elegant, efficient and successful, yet rather haphazard solution to this problem is offered by two seemingly disparate but de facto identical “theories”: radiative transfer theory that describes the photon exchange in terms of integro-differential equations, and Monte Carlo modeling that simulates the energy transport as a random flight of photons.

10.1 Radiative transfer

Radiative transfer relies heavily on the particle picture of light. One forgets about electromagnetic waves and regards photons as localized point-like particles that propagate along straight lines that are called rays in the context of geometrical optics. Occasionally, a photon bounces with a material particle and changes the direction of the flight. The only role of the wave optics is to provide rules for such photon-particle encounters, namely the scattering cross-section $\sigma_t$ and the phase function $p(\theta,\phi)$. The remaining part of the radiative transfer theory consists in the accurate accounting of rays and photon fluxes, under the strict observation of the law of energy conservation. The presumptions of Radiative Transfer can be formulated as follows: I) Light fields can be replaced by a manifold of geometrical rays. II) Nevertheless, scattering and absorption processes can be characterized by the Beer’s law and by the pre-averaged phase function $p(\theta,\phi)$ (recall Sec. 9.3). III) Light propagating along different rays does not interfere.

We shall now investigate if and how are those presumptions compatible with the electromagnetic theory of light. The key idea needed to justify the ray-preservation is to realize that a ray is the limiting case of a beam-wave (recall Sec. 2.3.1). Ray is a beam wave whose effective radius $a_b$ and effective solid angle $\Omega_b = 4\pi/k^2a_b^2$ have shrunk to zero, while the Rayleigh range $z_r = ka_b^2$ extends to infinity. The only way to reach this limit without violating the uncertainty principle is to shrink the wavelength $\lambda$ to zero. In this sense, geometrical optics is the short wavelength limit of wave optics [37]. Geometrical ray optics precludes all kind of interference. Fortunately, however, the transition into the geometrical limit is unnecessary. A beam appears as a ray for all practical purposes if the following conditions are fulfilled: i) The Rayleigh range $z_r$ (recall Fig. 2.2) is much larger than the size $L$ of the region of interest. ii) The effective radius $a_b$, that eventually determines the resolution, is much smaller than $L$. These conditions can be quite easily fulfilled with visible light in stellar and planetary atmospheres: a beam of a radius $a_b = 1\text{km}$ and Rayleigh range $z_r = 6 \times 10^9\text{km}$ (at $\lambda = 1\mu\text{m}$) will certainly appear like a ray when looked at from Proxima Centauri.

Having accepted beams in place of geometrical rays, we are able to analyze the scattering processes in a rigorous wave-optical way, as we have done in previous chapters. Thereby, one realizes that the conditions on the “beam-rays”, that are needed to justify presumption II, are actually more strict: We must require that $a_b \gg \lambda_C$, where $\lambda_C$ is the structural correlation length of the scattering system. On the other hand, in order to discern single scattering events we must require that $a_b \ll \ell_e$, where $\ell_e$ is the extinction length of the Beer’s law. Fortunately we may somewhat release the condition on the Rayleigh range: it is sufficient to require that $z_r \gg \ell_e$. All this can be certainly well fulfilled in the optics of atmospheres, but less clear is the situation in biological tissues (recall the concluding remarks in Sec. 9.4).

The third presumption of Radiative Transfer concerns the absence of interference of light waves propagating in different “beam-rays”. To justify this presumption, we must first reflect on the meaning of the phrase “absence of interference”. Recall from
Sec. 2.3.1 that any propagating light field can be expressed as the superposition of plane waves. Thereby, the total power carried by the light field is the sum of the partial powers carried by the individual partial waves\(^1\). Expressed in the language of mathematics, plane waves represent a complete set of orthogonal base vectors in the space of the solutions of Maxwell Equations. The construction of a beam wave, Eq. 2.29, is just one example. Since a beam-wave in Rayleigh range is a quasi-plane wave, one can expect that a propagating light field can be as well constructed as a superposition of orthogonal unit beam waves, at least locally and in a good approximation. Recalling the notion of the unit beam wave from Sec. 5.2, we write the resulting field as

\[
|E\rangle = \sum_b |b\rangle \mathcal{E}_b
\]

Since the orthogonal modes do not interfere, the power \(P_E = \langle E|E\rangle\) is the sum of powers carried in the individual modes:

\[
P_E = \sum_b |\mathcal{E}_b|^2
\]

Note that “absence of interference” does not mean that the light fields of the partial modes would not interfere locally, forming fringes or speckles wherever they overlap. The term “absence of interference” merely asserts the validity of Eq. 10.5: there are no cross-terms, i.e., no interference between the orthogonal modes. Note that the term “orthogonal” does not only imply the orthogonality of the polarization vectors \(\hat{e}_b\) or the propagation vectors \(\hat{s}_b\) of the base modes. Two beam waves \(|b_1\rangle\) and \(|b_2\rangle\) are orthogonal if \(\langle b_1|b_2\rangle = 0\). The scalar product \(\langle b_1|b_2\rangle\) is defined in Eq. 5.4. Using the paraxial approximation, Eq. 5.5, one can verify that two Gaussian beam waves \(|b_1\rangle\) and \(|b_2\rangle\) are nearly orthogonal as soon as the angle between their propagation vectors \(\hat{s}_1\) and \(\hat{s}_2\) exceeds a couple of the beams divergence angles \(\beta = 1/ka_b\). Two beams are also nearly orthogonal as soon as the shortest distance between their central rays exceeds a couple of effective beam radii \(a_b\). (Gaussian beams may be not a best set of orthogonal base vectors, but sufficient for the present plausibility argument.) The sketch in Fig. 10.1A illustrates a selected few of orthogonal “beam-rays” crossing a certain surface \(\Pi\). Note \(^1\)Since plane waves are delocalized, their individual power contributions are infinitesimally small.

![Figure 10.1](image_url)

A: B:

that each of the beams exhibits a small yet finite effective solid angle \(\Omega_b\), which is not easy to visualize, because \(\Omega_b\) is small and therefore the Rayleigh range \(z_r\) rather long.

Having understood the concept of the superposition of orthogonal modes, we are in the position to make the transition to Radiative Transfer plausible. Consider a planar surface of an area \(A\) that is much larger than the mode area \(\pi a_b^2\). This plane is irradiated
by the field $|E|$, that is a superposition of a large number of the orthogonal beam waves. For simplicity we assume that all modes have the same effective cross-section $\pi a_b^2$ and the same effective solid angle $\Omega_b$. The modes are distinguished by two parameters: the interchapter $\rho_b$ of their principal ray with the plane $\Pi$ and the propagation vector $\hat{s}_b$.

If one could record the irradiance in the $\Pi$ plane, one would observe a speckled pattern, illustrated in Fig. 10.1B. This pattern reflects local interference effects resulting from the superposition of the beam waves with a broad distribution of propagation vectors $\hat{s}_b$.

Since this distribution extends over the entire $4\pi$ solid angle, the average size of a speckle corresponds roughly to one wavelength $\lambda$. In Radiative Transfer we are not interested in such details, but only in length scales that are larger than $a_b$. Thus, we pre-average the pattern over an area of a couple of $\pi a^2_b$. This pattern reflects local interference effects resulting from the superposition of the beam waves with a broad distribution of propagation vectors $\hat{s}_b$.

Recall that in electrodynamics the transmission of power across a surface $\Sigma$ is expressed as

$$P = \int S \cdot \hat{n} d\Sigma$$

In the world of rays, $\Omega_b$ and $\pi a^2_b$ are regarded as infinitesimal quantities. Thus, we replace the summation by integration:

$$P = \int \int \int L(\rho, \hat{s}) \, \hat{s} \cdot \hat{n} \, d\Sigma$$

Recall that in electrodynamics the transmission of power across a surface $\Pi$ is expressed as

$$P = \int \int S \cdot \hat{n} d\rho = \int \int S \cdot \hat{n} d\rho = \int u \hat{s} \cdot \hat{n} \, \delta \rho^2,$$

where $S$ is the local Poynting vector, $u$ is the local energy density and $\hat{n}$ is the speed of energy transport. By the mode pre-averaging we arrived at the ray version of this relation, which can be expressed as

$$P = \int \frac{F \cdot \hat{n} d\rho}{\pi a^2_{b}} = \int \frac{\phi \hat{s} \cdot \hat{n} d\rho}{\pi a^2_{b}} = \int \frac{u \hat{s} \cdot \hat{n} \, \delta \rho^2}{\pi a^2_{b}}$$

where

$$F(r, t) = \int_{4\pi} L(r, t, \hat{s}) \hat{s} d\Omega, \quad \phi(r, t) = \int_{4\pi} L(r, t, \hat{s}) d\Omega, \quad u = \frac{\phi(r, t)}{v}$$

Here $F(r, t)$ is the pre-averaged energy flux density and $\phi(r, t)$ is the so-called fluence rate. In Eq. 10.9 $\bar{s} = F(r, t)/\phi(r, t)$ is the mean propagation vector. What remains to establish the Radiative Transfer is to invoke energy conservation, that can be expressed through the Equation of Continuity:

$$\frac{\partial \bar{u}}{\partial t} = - \nabla \cdot F - \text{sinks + sources} \quad (10.11)$$

In Sec. 11 we shall also reflect on global interferences of photon paths. Those are not apparent here, since Radiative Transfer is a local theory.
The last step is to recall that the symbol \( L(r, \hat{s}) \) actually represents a beam wave passing through \( r \) in the direction \( \hat{s} \). During the propagation, energy is removed from each beam wave by absorption and scattering, as expressed by the Beer’s law with the extinction coefficient \( \mu_e = \mu_a + \mu_s \). However, there is also energy input by scattering from the crossing beams from a direction \( \hat{s}' \) into the direction \( \hat{s} \). This applies independently for all beams, and therefore we can remove in Eq. 10.11 the \( \Omega \)-integration. Thus we write:

\[
\frac{1}{v} \frac{\partial L(r, \hat{s})}{\partial t} + \frac{\partial L(r, \hat{s})}{\partial s} = -\mu_e L(r, \hat{s}) + \mu_s \int_{4\pi} p(\hat{s}|\hat{s}') L(r, \hat{s}') d\Omega' + S_L \tag{10.12}
\]

Here we employed the identity \( \hat{s} \cdot \nabla L(r, \hat{s}) = \frac{\partial L(r, \hat{s})}{\partial s} \). The source term \( S_L \) is the reminder of the laser injecting somewhere energy into the system. Eq. 10.12 is the Equation of Radiative Transfer in the time dependent version. Note that it is quite similar to Boltzmann’s Equation for the transport of the molecules in a gas, simplified by the fact that all photons propagate with the same speed. The physical meaning of Eq. 10.12 can be easiest understood in the stationary case, when \( \partial L/\partial t = 0 \).

\[
\frac{\partial L(r, \hat{s})}{\partial s} = -\mu_e L(r, \hat{s}) + \mu_s \int_{4\pi} p(\hat{s}|\hat{s}') L(r, \hat{s}') d\Omega' + S_L \tag{10.13}
\]

On the right hand side, the first term represents the Beer’s law, i.e. energy removal from the beam wave over the infinitesimal length \( \partial s \). The second term accounts for photons that have been scattered within \( \partial s \) from all other modes \( \hat{s}' \) into the state \( \hat{s} \). Recall also the laser energy input \( S_L \).

The only difference between the scalar transport theory and its polarized version is that not only the propagation vector \( \hat{s} \) but also the polarization vector \( \hat{e} \) specify the photon state. Correspondingly, the “rules for photon-particle encounters” are a little bit more complex. In the scalar theory the rules consist in the phase function \( p(\hat{s}|\hat{s}') \) and the coefficients \( \mu_s, \mu_a \). In the reality, however, the photon exchange does not only depend on the propagation vectors \( \hat{s} \) and \( \hat{s}' \) of the involved photons, but also on their polarization states \( \hat{e} \) and \( \hat{e}' \). Since the polarization modes span a 2D complex space, it is advantageous to express the polarized rules in the matrix formalism of Mueller [14] or Jones [61].

The task would be now to specify the initial and boundary conditions, including the laser light source, and solve the integro-differential Eq. 10.12. This feels like a formidable amount of mathematical work, and therefore we give the radiative transfer up, just like we have given up the rigorous multiple scattering theory. Fortunately, there is an alternative approach, much more intuitive and conceptually simple than the abstract mathematics of radiative transfer.

### 10.2 Photon random flight and Monte Carlo simulations

Since we have given up the electromagnetic waves, we can as well understand the energy transport in the multiply scattering medium as a random flight of photons. In this picture photons are regarded as a kind of billiard balls that fly through the scattering medium with the transport speed \( v \), bouncing thereby randomly into some random obstacles. However naive, this picture is an exact representation of the scalar radiative transfer. It is somewhat surprising that Chandrasekhar does not refer to this simple model, despite having contributed significantly to the theory of random flights [62]. The most likely reason for his negligence is the lack of computing power: he did not consider the possibility, because in the forties still the only way of treating physical
problems was setting up and solving equations. Today we have the means to solve the transport equation directly, that is by Monte Carlo simulations of random flight. The great feature of the MC simulation is its dual nature: it is a flexible technique for solving transport equation with arbitrary shapes of the scattering medium and arbitrary phase functions, but it is at the same time a simple, intuitive and lucid model of the physical reality. With scalar photons, the MC simulation is exceedingly simple: one launches a photon into the tissue at a certain initial position $r_0$ and lets it fly a distance $d_0$ in the direction $\hat{s}_0$ until it bounces into a scatterer at $r_1$. This encounter changes the flight direction to $\hat{s}_1$ and the procedure is repeated. The photon pursues a highly irregular path through the medium, until it leaves the sample to vanish in space. Thereby Beer’s law is implemented in a probabilistic fashion, by throwing a dice: the flight distance between two interaction events is distributed according to the exponential distribution

$$p_d(d)\delta d = \frac{1}{\ell_e} e^{-d/\ell_e} \delta d \quad d \in \{0, \infty\} \quad (10.14)$$

where $\ell_e = \langle d \rangle = 1/\mu_e$ is the interaction mean free path, $\mu_e = \mu_s + \mu_a$ is the extinction coefficient (alias total attenuation coefficient). Thus, to generate an interaction event, one generates a random number out of the distribution 10.14. After the interaction, one more random number is needed to simulate the absorption: with the probability $\mu_s/\mu_e$ the photon proceeds its irregular flight, with the complementary probability $\mu_a/\mu_e$ it is absorbed. When absorption has occurred, one marks this event in a 3D matrix that will finally contain the energy deposition in the tissue. Otherwise the change of flight direction in the scattering process is simulated according to the chosen phase function $p(\hat{s}; \hat{s}')$. Thus, one generates two random numbers from the joint distribution $p_{\theta,\phi}(\theta, \phi) \sin(\theta) \delta \theta \delta \phi$, such that $\theta \in \{0, \pi\}$, $\phi \in \{0, 2\pi\}$. Note that the source term $\mu_s \int 4\pi p(\hat{s}; \hat{s}') L(r, \hat{s}') d\Omega$ does not explicitly enter MC simulations; it will be automatically realized after having generated a large number of random photon paths. Much more about such MC simulations can be read in Chap. 5. Finally we note that from random flight there is only a short step to the diffusion approximation. As shown, e.g., by Chandrasekhar [62], diffusion process is the limit of a random flight with an infinitesimally short scattering mean free path $\ell$, but with a correspondingly large number of steps $n$, so that $L_n = \text{const}$ while $1/\ell = \mu_s \to \infty$. (Thereby keep $\mu_a$ finite; absorbed photons do not propagate!) In this limit, the transfer equation, Eq. 10.12, can be replaced by the diffusion equation for the energy density $u(r, t)$, as discussed in detail in Chap. 6.
Monte Carlo simulations of disordered light propagation rely on the concept of photon path. This path approach involves serious approximations with respect to the exact electromagnetic theory, and one may ask why Monte Carlo is so successful. A rigorous transition from electrodynamics to radiative transfer and Monte Carlo simulation is a formidable task [63], well beyond the scope of the present contribution. However, we can at least try to make the approximations involved in this transition explicit. At the same time we take the opportunity for a brief summary of the concepts and tools which we have prepared in previous chapters. Unlike most polarized MC simulations previously published in the context of tissue optics [64, 65, 66, 67, 68, 69], we will not use Mueller formalism, but work on the amplitude level. We follow thus the path that is also taken by researchers interested more in fundamental than in applied aspects of light propagation [70, 71, 72]. The reasons have been mentioned earlier: i) single scattering usually does not depolarize and therefore there is no reason to employ Mueller, ii) Jones formalism works on the level of amplitudes and is therefore nearer to the underlying physics than Mueller’s averaged products of complex amplitude components, iii) we would like to keep the door open for future developments that would involve interferences of the photon paths.

11.1 Photon path

As we decided to work on the amplitude level, we find it convenient to employ a pedestrian version of Feynman’s concept of the path. We shall follow Feynman’s path rules, as given in [73], or, in the pedestrian version, in the enlightening booklet “QED - a strange theory of light and matter”[74]. (Feynman rules adapted for the present purpose can be found in [35], Chapter 20.) Photon path is a chain of localized scattering events connected by propagation. A photon is injected into a state $|l\rangle$ (this initial state is the impinging laser beam) and propagates to undergo the first scattering at $r_1$, which transfers the photon into a new state $|\pi_1\rangle$. Then it propagates until a second scattering event occurs at $r_2$, and so on, from $r_{i-1}$ to $r_i$, until the photon exits the scattering medium in a state $|\pi_p\rangle$ or is absorbed. There is a certain chance that an exiting photon ends up in a detector. In Feynman’s spirit we associate each path with a path amplitude $\mathcal{F}_p$ for the transition from an initial state $|l\rangle$ to a final state $|\pi_p\rangle$:

$$|\pi_p\rangle = \mathcal{F}_p |l\rangle$$  \hspace{1cm} (11.1)


In our case, the states \( |\pi_p\rangle \) are polarized beam waves characterized by a polarization vector \( \hat{e}_p \), propagation vector \( \hat{s}_p \) and a scalar complex amplitude \( \mathcal{E}_p \). Working with a succession of transverse coordinate systems, the path amplitude \( F_p \) can be represented by a \( 2 \times 2 \) Jones matrix and the states \( |\pi_p\rangle \) or \( |l\rangle \) as Jones vectors. However, Feynman’s approach is quite general, and therefore we keep using the general bra ket notation.

The path probability is the probability that a photon injected into the state \( |l\rangle \) will make it into the state \( |\pi_p\rangle \). The path probability is defined as:

\[
\mathcal{P}_p = \frac{\langle \pi_p | \pi_p \rangle}{\langle l | l \rangle} = \frac{\langle l | F_p^\dagger F_p | l \rangle}{\langle l | l \rangle} = \langle l | F_p^\dagger F_p | l \rangle.
\] (11.2)

On the right hand side we recall that the initial state is normalized, i.e. \( \langle l | l \rangle = 1 \). The path amplitude is the product of segment amplitudes, i.e., the photon state after the last scattering event is represented by

\[
|\pi_p\rangle = F_p |l\rangle = \Pi_{i=1}^p F_i |l\rangle,
\] (11.3)

where the segment amplitude \( F_i \) is the amplitude for the transition of a photon from a state \( |\pi_{i-1}\rangle \) to a state \( |\pi_i\rangle \), such that \( |\pi_i\rangle = F_i |\pi_{i-1}\rangle \). Obviously, the path probability \( \mathcal{P}_p \) can be expanded into a product of segment probabilities \( \mathcal{P}_i \). We only indicate the first step of the expansion:

\[
\mathcal{P}_p = \frac{\langle \pi_p | \pi_p \rangle}{\langle l | l \rangle} = \frac{\langle \pi_{p-1} | F_p | \pi_{p-1} \rangle}{\langle \pi_{p-1} | \pi_{p-1} \rangle} \frac{\langle \pi_{p-1} | \pi_{p-1} \rangle}{\langle l | l \rangle} = \cdots = \Pi_{i=1}^p \mathcal{P}_i.
\] (11.4)

Here \( \mathcal{P}_i \) is the probability that a photon that entered the photon state \( |\pi_{i-1}\rangle \) at \( r_{i-1} \) will make it to \( r_i \) and there becomes scattered into a state \( |\pi_i\rangle \):

\[
\mathcal{P}_i = \frac{\langle \pi_i | \pi_i \rangle}{\langle \pi_{i-1} | \pi_{i-1} \rangle} = \frac{\langle \pi_{i-1} | F_i^\dagger F_i | \pi_{i-1} \rangle}{\langle \pi_{i-1} | \pi_{i-1} \rangle}.
\] (11.5)

Notice that in the present context “probability” is to be understood in the quantum sense, as the complex square of an amplitude. We are now dealing with pure states and the fluctuations in the scattering medium are not yet considered.

A real experiment involves a receiver and a detector. We are interested in the normalized expected signal, i.e., in the detection probability \( \mathcal{P}_d = N_d / N \), where \( N_d \) is the number of photons that are likely to be detected, when \( N \) photons were injected into the system. For simplicity we assume an ideal detector, which registers with a probability 1 each photon that enters into the receiver. The receiver is a polarized single-mode beam filter (Sec. 5), which projects each of the incoming states \( |\pi_p\rangle \) on the receiver mode \( |o\rangle \) (recall that \( \langle o | o \rangle = 1 \)). The received state \( |\pi_d\rangle = |o\rangle \mathcal{E}_d \) is “filled” with contributions from all paths:

\[
|\pi_d\rangle = |o\rangle \mathcal{E}_d = |o\rangle \sum_p \langle o | F_p | l \rangle.
\] (11.6)

The receiver mode \( |o\rangle \) and the initial state \( |l\rangle \) are normalized, so that we write the detection probability as

\[
\mathcal{P}_d = \frac{\langle \pi_d | \pi_d \rangle}{\langle l | l \rangle} = \sum_{p=1}^{N_p} \sum_{q=1}^{N_p} \langle l | F_p^\dagger | o \rangle \langle o | F_q | l \rangle.
\] (11.7)

Note that there are two types of terms in the double sum: The \( N_p \) terms with \( p = q \) give the incoherent sum of the detection probabilities of the individual path \( P_{dp} = \).
The cross-terms with $p \neq q$ express the interference of paths. In Sec. 10.1 we did not consider this type of interference, because in Radiative Transfer the paths are not explicitly visible; all one has is the local transfer equation, Eq. 10.12. However, we discussed the local interference that leads to the speckle pattern shown in Fig. 10.1B. The same pattern would appear on an area detector positioned in place of the single mode receiver. Integrating the signal over the detector area $A$, the speckles will average out: because a multimode detector receives a large number of states, one can invoke in Eq. 11.7 the closure relation $\sum_o |o\rangle \langle o| \approx I$ and assume that the detected signal is the sum of probabilities of paths terminating in the detector area:

$$P_A \sim \sum_{A,p}^{N_p} \langle l| F_p^\dagger F_p |l \rangle = \sum_{A,p}^{N_p} P_p.$$  (11.8)

This kind of detection is employed in the classical scalar MC simulations. However, the role of interference and fluctuations is precisely what we would like to address in future work. Therefore we use a single-mode detector and the general expression for the detected signal, Eq. 11.7. Thus, we must explicitly carry out the last step of the path analysis, namely the time averaging. All scattering media of practical interest exhibit thermal fluctuations, which keep rearranging their microscopic structure. Thus, the amplitude of a path through a fixed set of scattering locations $\{r_i\}$ fluctuates in time, $F_p = F_p(t)$. The measured detection probability is the time average:

$$\overline{P_d} = \frac{1}{T} \int_0^T P_d(t) \, dt = \sum_{p} \sum_{q} \langle l| F_p^\dagger(t) |o \rangle \langle o| F_q(t) |l \rangle$$  (11.9)

Here the cross terms represent paths correlations. Strictly speaking, only if all paths are uncorrelated, then we are allowed to write the detected signal as the sum of the detection probabilities of the individual paths:

$$\overline{P_d} \approx \sum_{p} \langle |o| F_p(t) |l \rangle \rangle^2$$  (11.10)

To justify the decorrelation we must start filling the bra-kets with their contents.

**11.2 Initial photon state, launching a photon**

A realistic initial photon state is a collimated or moderately focused laser beam. We model the beam as a paraxial superposition of plane waves which span a certain range of the wave vectors around a principle wave vector $k_l$ and which are synchronized in the focal point $r_l$:

$$|l \rangle = |\hat{e}_l, k_l, r_l \rangle \equiv \sqrt{\frac{2}{c \epsilon_m}} X(r; r_l, k_l) \hat{e}_l e^{i k_l \cdot (r - r_l)}.$$  (11.11)

The beam profile $X(r; r_l, k_l)$ is for example a Gaussian $X(r) = 1/a(z) \exp(-\rho^2/2a(z)^2)$, where $a(z) = a_l \sqrt{1 + z^2/z_r^2}$ is the effective beam radius, that varies with the distance $z$ from the beam focus, but within the Rayleigh range $z < z_r = ka_l^2$ it is approximately constant. We assume that the length $z_r$ is much larger than the extinction length $\ell_e = 1/\mu_e$, so that we can regard the beam as a quasi plane wave in the relevant range of propagation. Taking a typical $\ell_e \approx 30\mu m$, the condition $z_r \gg \ell_e$ is already fulfilled with a beam as narrow as $a_l = 10\mu m$. 

81
Figure 11.1: Multiple scattering model: photon propagation through a system of paths of beam-rays connecting virtual sample-cells.

Into the normalized beam $|l\rangle$ we start launching photons. The squared profile $|X(\rho)|^2$ can be interpreted as the probability density for finding a photon in a certain distance $\rho$ from the beam axis. Thus, we choose a unique initial propagation vector $s_0$ and launch the photons with an offset $\rho$ with respect to the beam axis $\hat{z}$ so that $\rho_x$ and $\rho_y$ are random numbers from the distribution $X^2(\rho_x, \rho_y)$. A rigorous treatment of launching through a strongly focused laser beam or using a fiber-optic optode would require more thinking. As a first order approximation one can adopt the approach of radiative transfer, replacing the light fields with polarized photon streams.

### 11.3 Scattered photon states, beams and paths

The scattered fields are more complicated. They result from the coupling of dipole fields of a large number of elementary microscopic scattering processes. In other words, scattering involves a large number of elementary photon paths. Such coupled dipole approach has been successfully used to calculate the scattering from small particles [60], but it is entirely inadequate for light propagation in macroscopic disordered media. The basic idea of the MC simulation is essentially to “preprocess” the maze of elementary paths into manageable sets of path bundles. In most MC simulations one implicitly assumes that the scattering medium consists of isolated spherical particles. The scattering events are thought to be localized in their centers and a photon propagates along ray-segments connecting the individual particles. We shall keep this picture, but replace in mind the tangible Mie-spheres with the ghost quasi-particles from Sec. 9.2.1, i.e. with fuzzily localized regions in space surrounding each scattering event. Moreover, we replace the rays by orthogonal beam waves, in the same spirit as we have already done in Sec. 10.1, while “deriving” the Equation of Radiative Transfer. In other words, we decompose the spherical wave radiating from a quasi-particle into a broad range of wave vectors, into a more or less complete set of orthogonal partial beam waves, as illustrated in Fig. 11.1. Each of the “beam-rays” radiates into a small solid angle $\delta\Omega$. The procedure is analogous to the rigorous expansion of a radiating dipole wave into plane waves [75], but we prefer quasi-plane beam waves, in order to preserve the localization of the path. Admittedly, the proposed kind of decomposition is rather fuzzy. Nevertheless, if the procedure succeeds, then we managed to provide a sound electromagnetic ground to Monte Carlo simulations in multiply scattering media.

The focal point of each beam is the location $r_i$ of the scattering event $i$ and the
principle wave vector \( \mathbf{k}_i \) is directed along the path segment vector \( \mathbf{r}_{i+1} - \mathbf{r}_i \). These scattered beams may be thought to have a certain finite waist radius \( \delta d \) (equivalent to the Gaussian beam radius \( a_0 \)) and a corresponding solid angle \( \delta \Omega = 4\pi/(k\delta d)^2 \). In each beam we collected a bundle of elementary photon paths that are synchronized in \( \mathbf{r}_i \) and that propagate in the solid angle \( \delta \Omega \). In a scattering event the photon “chooses” one of such path bundles, as illustrated in Fig. 11.1. One may imagine the scattering system as being subdivided into “sample cells”, formed at the interchapters of the partial scattered beam waves. In between the cells the beam wave propagates in a macroscopically homogeneous effective medium, the only reminder of the microscopic inhomogeneity being the extinction of the beam wave due to scattering or absorption. The cell \( C_i \) is centered in the \( i \)-th scattering event, that is in the focal point of the beam wave \( |\pi_i \rangle \). In each realization of path a new subset of such cells is created. The volume \( V_i \) of each cell is given by the waist radius of the scattered beam waves, i.e. \( V_i \approx \pi(\delta d)^3 \). As we shall see later, the exact size does not matter. However, there are certain restrictions: the waist radius \( \delta d \) should be larger or at least comparable with the structural correlation length \( \lambda_s \), but much smaller than the mean distance \( \ell \) between the scattering events, so that only single scattering is likely to occur in each cell. In addition, the Rayleigh range \( z_r = k(\delta d)^2 \) of the partial beam wave should be large compared with \( \ell \), i.e., \( k(\delta d)^2 > \ell \). At this point we may summarize the requirements concerning the length scales:

\[
1 \ll k\ell \quad (k\lambda_s)^2 \ll k\ell < (k\delta d)^2 \quad k\lambda_s \ll k\delta d \ll k\ell \quad (11.12)
\]

If these relations can’t be fulfilled, then MC simulation likely fails to describe the reality.

### 11.4 Averages and correlations

The last but not least fundamental approximation involved in modeling the photon random flight is path decorrelation which results from the decorrelation of individual segments. A path segment consists of the propagation of the photon from \( \mathbf{r}_{i-1} \) to \( \mathbf{r}_i \) and of the scattering event at \( \mathbf{r}_i \). Having accepted the photon path picture in Fig. 11.1, we can immediately apply the results that we developed until chapter 9.3. The segment amplitude matrix \( \mathcal{F}_i \) can be concisely written as

\[
\mathcal{F}_i(t) = \mathcal{A}_i(t) \Pi^d(\theta_i) \mathbf{R}_2(\phi_i) \cdot \mathbf{J} e^{ikm d} e^{-\mu e d/2},
\]

In the present chapter we disregard the propagation that is included in the phasors and in the Jones matrix \( \mathbf{J} \). We also disregard the matrix \( \Pi^d(\theta_i) \mathbf{R}_2(\phi_i) \), which concerns only the polarization. We assume, as we discussed in Sec. 9.3, that the dielectric response does not fluctuate, so that an incoming pure polarization state remains a pure state upon scattering. The only fluctuating quantity is the scalar amplitude

\[
\mathcal{A}_i(t) = \frac{i k_m}{2} \int X_{i-1}(\mathbf{r}) X_i(\mathbf{r}) \delta \chi(\mathbf{r}, t) e^{i \mathbf{q} \cdot \mathbf{r}} d^3 r.
\]

Here \( \delta \chi(\mathbf{r}, t) \) represents susceptibility fluctuations, \( \mathbf{q} \) is the scattering vector and \( X_{i-1}(\mathbf{r}) \), \( X_i(\mathbf{r}) \) are the profiles of the incoming and outgoing beams. Note that \( \mathcal{A}_i(t) = 0 \), since the time average of \( \delta \chi(\mathbf{r}, t) \) is zero by definition. To investigate the decorrelation we return to Eq. 11.9 which we rewrite explicitly in terms of the amplitudes \( \mathcal{F}_{pq} \) of the \( i \)-th segment in the \( p \)-th path:

\[
\overline{\mathcal{F}}_d = \sum_p \sum_q \langle \ell | \Pi^p_{i=1} \mathcal{F}^\dagger | o \rangle \langle o | \Pi^q_{j=1} \mathcal{F}^\dagger | \ell \rangle \quad (11.15)
\]
The time averaging concerns the products $\mathcal{A}_{pi}^\dagger(q_i, r_i, t) \mathcal{A}_{qj}(q_j, r_j, t)$. Recall the meaning of “uncorrelated”: the average of the product of two independent, or *uncorrelated*, random variables can be factored into the product of their averages. The decorrelation approximation is footed on the following fact: two scalar segment amplitudes $\mathcal{A}_{pi}(q_i, r_i, t)$ and $\mathcal{A}_{qj}(q_j, r_j, t)$ are uncorrelated unless they belong to the same scattering volume (the same quasi-particle) at $r_i$ and the same scattering vector $q_i$. More precisely, as soon as $|r_i - r_j| > \delta d$ and/or $|q_i - q_j| > 1/\delta d$ then

$$
\mathcal{A}_{pi}^\dagger(q_i, r_i, t) \mathcal{A}_{qj}(q_j, r_j, t) = \mathcal{A}_{\pi}^\dagger(q_i, r_i, t) \mathcal{A}_{q_\pi}(q_j, r_j, t) = 0 \quad (11.16)
$$

The lack of correlation between two non-overlapping volumes is obvious. The $q_i q_j$-decorrelation requires a little bit work [76], but finally one finds that two amplitudes decorrelate as soon as the difference of the scattering angles is larger than the divergence of $1/ka$ of the involved beam wave. We conclude: two paths are uncorrelated if there is at least one distinct segment satisfying Eq. 11.16. Such a situation is illustrated in Fig. 11.2A. Uncorrelated segments can be found in a large majority of paths. Therefore we can neglect in Eq. 11.16 the cross-terms with $p \neq q$ and use the decorrelation approximation, Eqs. 11.10.

![Figure 11.2: Path decorrelation. A: Paths $p$ and $q$ are decorrelated because of the distinct branching. B: In the path $p$ there is a loop with one correlated segment. We neglect such loops.](image)

It is obvious, that the lack of inter-segment correlation justifies also using the pre-averaged version $\overline{T}$ of the transition matrix, recall Eq. 9.31. Neglecting the rare path loops indicated in Fig. 11.2B, we can write the time averaged path probability (recall Eq. 11.2) as the product of time averaged segment probabilities:

$$
\overline{P}_p = \prod_{i=1}^{n_p} \overline{P}_i = \prod_{i=1}^{n_p} \overline{P}_i.
$$

(11.17)

Recall, however, the restriction at the end of Sec. 9.3: if the quasi-particles consist in an ensemble of anisotropic scatterers with randomly fluctuating orientation of the anisotropy axis $\hat{a}$, we would probably have to resort to Mueller formalism.

The de-correlation approximation is not entirely obvious in the case of a medium with frozen microscopic heterogeneity, such as biological tissues most likely do exhibit. Lacking a better theory we content ourselves with the assumption that even in such non-ergodic case we are allowed to use an ensemble pre-averaged segment amplitude. In any case, we are likely to miss certain effects of interference neighboring paths. For example, the coherent back scattering that results from interference of counter-propagating but otherwise identical paths [71] is excluded by the decorrelation approximation.
11.5 Polarized segment amplitude and probability

Having simplified the problem by decorrelation of the paths and segments, we may focus on a single segment, in particular on the effect of polarization. The geometry of the segment is illustrated in Fig. 11.3. Notice that the angles $\phi_i$ and $\theta_i$ are Euler angles (as defined in [77]) of rotations relative to the coordinate system $i-1$. We take it for granted that the transformations from a local coordinate system to the laboratory system have been carried on together with the evolution of the photon path. In other words, we assume that we know a 3D rotation matrix $R_i$ constructed from the succession of $\phi$ and $\theta$ rotations:

$$R_i = \prod_{j=0}^{i-1} R_{\phi}(\theta_j) R_{\phi}(\phi_j)$$  (11.18)

The element $j = 0$ includes the transformation concerning the initial state $|l\rangle$, i.e. the aiming and polarization alignment of the laser beam. Recall that the segment amplitude matrix $F_i$ can be concisely written as

$$F_i(t) = A(t) \Pi_d(\theta_i) R_2(\phi_i) \cdot J e^{i k_m d} e^{-\mu_e d/2}$$  (11.19)

The dot $\cdot$ separates the factors that concern propagation of the photon from those that concern scattering. We discuss first the propagation.

11.5.1 Propagation

The factor $\exp(i k_m d)$ is the phasor associated with the optical path $k_m d$, where $d = |r_i - r_{i-1}|$. The subscript $m$ in $k_m$ is the reminder of the assumption that the beam propagates in an effective medium of refractive index $n_m$. The factor $\exp(-\mu_e d/2)$ describes in a crudely simplified fashion the removal of photons from the partial beam $|\pi_{i-1}\rangle$ by scattering and absorption. The extinction coefficient $\mu_e = \mu_s + \mu_a$ includes both processes. The evolution of the polarization state during the propagation is effected by the Jones matrix $J$, such that $J^\dagger J = I$, where $I$ is the 2D identity matrix ($J$ is a unitary matrix). In a plain isotropic medium $J = I$, and almost equally easy is an isotropic chiral medium. As discussed in chapter 2.2.2, a chiral medium causes a phase shift between the two circular polarizations, and therefore a rotation of linear polarization. The chiral propagation matrix is given by Eq. 3.25:

$$J = \begin{pmatrix} \cos(\delta k d) + \sin(\delta k d) \\ -\sin(\delta k d) \cos(\delta k d) \end{pmatrix}$$  (11.20)
where $\delta k = k_L - k_R$. In $\mathbf{J}$ one could easily include also circular dichroism, i.e. different attenuation for left and right circularly polarized beam waves.

Considerably more challenging is the case of anisotropy of the dielectric response, because in the course of their random flight the photons propagate in arbitrary angles with respect to the optical axis $\hat{a}$. This is much unlike the propagation through a well aligned retarder plate in Sec. 3. In chapter 2.2.1 we discussed the uniaxial anisotropy with arbitrary aligned axis, so that we now appreciate the difficulties. A major problem is the birefringence, i.e. the splitting of the beam wave in an ordinary and an extraordinary beam. The photon has two choices of $\hat{s}$ to propagate. Thus, in the MC simulation we should throw a dice to decide which way to take. But in this way we would miss the interference between the two polarization states and therefore also the possibility of producing circular or elliptical polarization out of linear (recall Sec. 3.4.3). Fortunately, the anisotropy of the biological tissue is not too large, so that the expected angles between the propagation vectors of the ordinary and extraordinary beam (Eq. 2.25) are perhaps smaller than the divergence angle $1/k\delta d$ of the beams. Thus, a good approximation may be to give up the birefringence and propagate the ordinary and extraordinary photon in the same beam wave. This approximation was used, for example, by Gangnus et al. [72] in their polarized MC simulation. Even so, the inclusion of anisotropy is rather demanding, since the propagation matrix must be newly constructed in each segment of the random flight.

### 11.5.2 Scattering

In the scattering factor $\mathcal{A}(t) \mathbf{\Pi}^d \mathbf{R}_2$ the first operator to be applied is the 2D rotation matrix $\mathbf{R}_2(\phi_i)$ which actuates the rotation from the coordinate system $i-1$ into the intermediate coordinate system $I$ in the scattering plane of the $i$-th scattering event; see Fig. 11.3 and recall Sec. 7.1 for a more detailed explanation. The matrix $\mathbf{\Pi}^d$ is the dipole polarization matrix, as introduced in Eq. 7.42. For convenience we recall that

$$\mathbf{R}_2(\phi) = \begin{pmatrix} \cos \phi & \sin \phi \\ -\sin(\phi) & \cos(\phi) \end{pmatrix}$$

Finally, $\mathcal{A}(t)$ represents the fluctuating scalar scattering amplitude. In the continuum model we write

$$\mathcal{A}(t) = \frac{ik_m}{2} \int X_{i-1}(\mathbf{r}) X_i(\mathbf{r}) \delta(\mathbf{r}, t) e^{i\mathbf{q} \cdot \mathbf{r}} d^3r,$$

(11.22)

Since no depolarization occurs in the single scattering process (recall Sec. 9.3 and 11.4), we simplify the matters by replacing $\mathcal{A}(t)$ with its pre-averaged version that is included in the time averaged transition matrix $\mathbf{T}$ in Eq. 9.31. Furthermore we recall the cell model discussed in chapter 11.5 and insert in Eq. 9.31 the following symbols: $V_{IOR} \approx (\pi \delta d)^3$, $\pi a_i^2 = \pi \delta d^2$ and $\Omega_o = \delta \Omega = 4\pi/(k\delta d)^2$. Thus, we replace Eq. 11.13 by

$$\mathcal{F}_i = i\sqrt{\mu_s \delta d \delta \Omega} \hat{S} \mathbf{R}_2(\phi_i) \cdot \mathbf{J} e^{ik_m d} e^{-\mu_d d/2}.$$

(11.23)

Here $\hat{S}$ is the normalized amplitude scattering matrix. In Born approximation:

$$\hat{S} = \frac{\sqrt{F(q_i)}}{\Phi} \sqrt{\frac{3}{8\pi}} \mathbf{\Pi}^d(\theta_i)$$

(11.24)

where $F(q_i) = F(\theta_i, \phi_i)$ is an experimentally determined or modeled structure factor and $\Phi$ is the already well known normalization constant, namely the integral of $F(\theta, \phi)$ weighted with the dipole phase function from Eq. 7.44:

$$\Phi = \int F(\theta, \phi) p_d(\theta, \phi) \sin(\theta) d^3d$$

(11.25)
Recall that in a structurally isotropic medium (in statistical sense), the structure factor \( F(q) \) depends only on \( |q| = k_n \sqrt{2 - 2\cos(\theta)} \).

At this point we would like to recall, that unlike all previously published polarized MC simulations we work here with polarized RDGB scattering, instead of scattering from Mie-spheres. Of course, a dilute suspension of monodisperse Mie particles is an excellent test system, good for a comparison of a simulation with a well controlled experiment. If desired, then one can employ the normalized \( \hat{S} \)-matrix of a Mie particle instead of \( \hat{S} \) from Eq. 11.24; recall the discussion in Sec. 9.3. However, we prefer to regard the tissue as a fluctuating continuum that can be treated in RDGB approximation. This allows to factor the scattering amplitude into the structure amplitude \( \sqrt{F(\theta)} \) that reflects the shape of the fluctuation correlation function, i.e. the mean structure of the medium, and in the matrix \( \Pi \) that accounts for the polarization. This factoring is a convenient starting point for future developments that will include anisotropy. Recall that there are two kinds of anisotropy to be encountered in the tissue: i) Anisotropy of structure, which concerns the structure factor \( F(q) \) and ii) anisotropy of dielectric response, that concerns the polarization matrix \( \Pi \). A generalization that would include both kinds of anisotropy into Monte Carlo simulations would be straightforward. However, anisotropic \( F \) and \( \Pi \) depend on the orientation of the local coordinate system with respect to the laboratory frame, which keeps changing in every step of the path. The implementation of anisotropies is an exercise in geometry that requires a considerable portion of patience and concentration.

### 11.5.3 Segment probability

It remains to calculate the segment probability \( P_i \), i.e. the probability that a photon that entered the photon state \( |\pi_{i-1}\rangle \) at \( r_{i-1} \) will make it to \( r_i \) and there becomes scattered into a state \( |\pi_i\rangle \) where it propagates in the direction \( \hat{s}_i \). We introduce the segment amplitude \( \mathcal{F}_i \) from Eq. 11.23 into Eq. 11.5 and rearrange the terms to obtain

\[
\mathcal{F}_i(\theta, \phi | \hat{e}_{i-1}) = p_i(\theta, \phi | \hat{e}_{i-1}) \delta \Omega \left[ \frac{\mu_s}{\mu_e} \cdot e^{-\mu_e d_{i}} \right] \delta d
\]

Equation 11.26 deserves to be framed, since it represents the fundamental of polarized Monte Carlo simulations. Note that neither \( \delta d \) nor \( \delta \Omega \) are used in the actual MC simulation. They only occur as concepts that allowed to derive Eq. 11.26 more or less rigorously, instead of simply assuming it.

It is important to realize that the segment probability \( P_i(\theta, \phi | \hat{e}_{i-1}) \) is the product of three probabilities that are separated by \( \cdot \) dot. Reading from right to left, the first factor is the probabilistic representation of the Beer’s law: \( \mu_e \exp(-\mu_e d_i) \delta d \) is the probability that a photon survives the propagation over a distance \( d_i = |r_i - r_{i-1}| \) from \( r_{i-1} \) to \( r_i \), but there it interacts with the matter within a small volume of size \( \delta d \). The second factor \( \mu_s / \mu_e = \mu_s / (\mu_s + \mu_a) \) is the probability that the interaction event is a scattering event, not an absorption. With the probability \( \mu_s / \mu_e \) the photon continues its path. With the probability \( 1 - \mu_s / \mu_e \) the path is terminated and the energy of the photon deposited in the matter. The third factor represents the phase function, i.e., the angular distribution of the scattered photons into different outcoming beam waves with propagation vectors \( \hat{s}_i \equiv (\phi_i, \theta_i) \) and solid angle \( \delta \Omega \):

\[
p_i(\theta, \phi | \hat{e}_{i-1}) = \frac{\langle \pi_{i-1} | \mathbf{R}_2^{-1} \hat{S} \hat{S} \mathbf{R}_2 | \pi_{i-1} \rangle}{\langle \pi_{i-1} | \pi_{i-1} \rangle} = \frac{F(\theta, \phi)}{8\pi} \frac{\langle \pi_{i-1} | \mathbf{R}_2^{-1} \Pi \Pi \mathbf{R}_2 | \pi_{i-1} \rangle}{\langle \pi_{i-1} | \pi_{i-1} \rangle},
\]

\[\tag{11.27}\]
where $|\pi_{i-1}\rangle = \exp(-\mu \epsilon d_i/2)J|\pi_{i-1}\rangle$. In Eq. 11.27 one recognizes the polarization factor of a dipole: $\langle \pi'_{i-1}|R_2^{-1}\Pi^d\Pi^dR_2|\pi_{i-1}\rangle/\langle \pi'_{i-1}|\pi_{i-1}\rangle$. We refrain from recalling the explicit result, since in a simulation algorithm one would anyway operate with matrices and vectors instead with lengthy scalar formulas. For the structure factor $F(\theta, \phi)$ we shall insert the proposed general expression from Eq. 9.37. One must also calculate the constant $\Phi$, Eq. 11.25. In the special case of HG structure factor, this work has been already done by Liu and Weng [78], who also provide a re-interpretation of the asymmetry parameter $g$ that takes polarization into account. These authors were probably the first to combine HG phase function with the dipole scattering matrix; naturally in the context of light propagation in the atmosphere.

11.6 Detecting a photon

Photon detection is an important stage of the simulation process since only now the photon is actually realized, re-emerging, so to speak, from the maze of the virtual path system into the physical reality. We would like to model the reality as closely as possible and convenient, and therefore we consider a realistic experimental setup for polarization imaging that is outlined in Fig. 11.4. Note that we are employing a telecentric imaging system so that the sample is everywhere imaged from the direction that is perpendicular to the sample surface. In the collimated range between the two lenses there is ample space to insert the optical elements necessary for polarization analysis. Thus, each of the imager channels is a combination of a beam filter (Sec. 5) and a polarization filter (Sec. 3.4). Only one observation beam is indicated by shading. Recall from chapter 2.3.1 that the polarization state remains preserved upon moderate focusing by a good lens, but take care to keep the angle of all observation channels low with respect to the optical axis. Not shown in Fig. 11.4 is the laser beam. A clever scheme of beam launching is crucial, since we should avoid specular reflection. Examples of practical realizations of polarization imaging can be found in Ref. [2].

With this setup in mind we must form an image from the simulated polarized photons emerging from the tissue. There are two ways how to understand image formation in terms of “physical optics”, i.e. optics of electromagnetic waves, as opposed to the optics of geometric rays. The easier approach is due to Rayleigh. One imagines in the object plane infinitesimal light sources, small particles that scatter the illuminating light. Each of the particles radiates a spherical wave, from which a beam wave is filtered out by the objective aperture and re-focused into the geometrical image point in the image plane. The image is formed by the superposition of the amplitudes of many such beams. Rayleigh’s approach is simple and intuitive, the only problem is that in our case it is difficult to discern any small particles in the object plane. We are dealing with a random superposition of waves that are radiated from the depth of the tissue. Thus, we must consider an alternative approach that was proposed by Abbe. Abbe forms the image out of waves only, without needing light sources in the object plane.
He regards the object as kind of gratting which modulates the phase and amplitude of the transmitted illuminating light. The resulting “complex disturbance” in the object plane is then propagated through the imaging optics according to Huygens-Fresnel principle and Kirchhoff-Fresnel theory of diffraction. To treat polarization, one should use the rigorous vectorial theory of diffraction, see e.g. [37].

Of course, the two approaches are fully equivalent. Interestingly, they correspond to our two representations of the scalar product of the observation mode with the observed field: whereas Rayleigh employs the source representation (Eq. 5.6), Abbe prefers the surface representation (Eq. 5.4) or rather the approximate version, Eq. 5.5. Note that using the virtual observation beam greatly simplifies Abbe’s approach: instead of propagating through the imaging optics a complicated superposition of waves, we propagate a well behaved selected beam. The idea is sketched in Figs. 11.4 and 11.5. We associate each pixel of the imager with a single-mode receiver beam. This is of course an approximation, since CCD pixels are neither single-mode fibers (not yet) nor points. But the approximation is good, if one takes care to match the pixel pitch with the resolution, which is given by the waist radius of the focused receiver beam in the object plane. Mathematically, this is expressed with our usual equation for the observation mode in the Rayleigh range:

$$|o\rangle \equiv \hat{E}_o^o(r) = \sqrt{\frac{2}{c\varepsilon_o}} X_o (r - r_o) \hat{e}_o e^{i k \hat{s}_o \cdot (r - r_o)}$$  (11.28)

where $X(r - r_o)$ is the amplitude profile of the beam focused in the object point $r_o$. The propagation vector $\hat{s}_o$ defines the alignment of the receiver. We chose $\hat{s}_o = -\hat{z}_L$, where $\hat{z}_L$ defines the positive $z$-direction in the laboratory coordinate system. Finally $\hat{e}_o$ represents one of the two unit vectors $\hat{x}_o$ and $\hat{y}_o$ that specify the linearly polarized base of the observation beam. A reasonable choice is $\hat{x}_o = \hat{x}_L$ and $\hat{y}_o = -\hat{y}_L$.

Knowing the light field $E^s(r)$ as it emerges from the sample in the plane Π, one can calculate, for each pixel, the projection amplitude as the overlap integral (Eq. 5.5) of $E^s(r)$ with the pixels receiver mode from Eq. 11.28:

$$\langle o|E^s \rangle \approx \frac{c \varepsilon_o}{4} \int_\Pi \hat{E}_o^o(\rho - \rho_o) \cdot E^s(\rho) \cdot [\hat{s}_s(\rho) + \hat{s}_o] \cdot \hat{n} d^2 \rho.$$  (11.29)

Here $\rho$ represents a 2D vector in the Π-plane. Recall from previous chapters that we regard the emerging field as being formed by the superposition of partial beam waves of the individual paths. Since the paths are assumed to be decorrelated, we can detect them individually, inserting the beam wave $E^p(\rho)$ of a path $p$ in place of $E^s(\rho)$ in Eq. 11.29. Only two paths are illustrated in the sketch in Fig. 11.5. Note that the paths exit the sample from a medium of refractive index $n_m$ into air. Thus, before
considering the detection, we would have to return to Sec. 3.5.1 in order to implement Fresnel’s refraction and reflection in the Jones matrix for the transmission across the (corrugated) interface. Here we only consider the Jones matrix for the transition from the transmitted beam-ray of the path \( p \) into the observation mode. Recall from Sec. 7.2 that

\[
\left( \begin{array}{c} E'_x \\ E'_y \end{array} \right) = \left( \begin{array}{c} \langle o_x | E_p \rangle \\ \langle o_y | E_p \rangle \end{array} \right) = T_{po} \left( \begin{array}{c} E'_o \\ E'_o \end{array} \right) \tag{11.30}
\]

Using the scalar product from Eq. 11.29, one obtains

\[
T_{po} = B_o(\rho_p, \hat{s}_p) \left( \begin{array}{cc} \hat{s}_o \cdot \hat{x}_p & \hat{s}_o \cdot \hat{y}_p \\ \hat{y}_o \cdot \hat{x}_p & \hat{y}_o \cdot \hat{y}_p \end{array} \right) \tag{11.31}
\]

Note that in order to calculate the polarization transfer matrix, all polarization unit vectors must be given in the same coordinate system, preferably in the lab system. Fortunately, we kept updating the 3D rotation matrix \( R \) during the evolution of the path (and in Fresnel refraction at the interface) so that now we can apply the operator \( R_p^{-1} \) to the photon polarization vector expressed in the last transverse coordinate system.

The scalar factor \( B_o(\rho_p, \hat{s}_p) \) is the overlap integral of the beam-ray \( p \) with the observation beam:

\[
B_o(\rho_p, \hat{s}_p) \approx \frac{[\hat{s}_o + \hat{s}_o] \cdot \hat{n}}{2} \int_{\Pi} X_o(\rho - \rho_o) X_p(\rho - \rho_p) e^{i q \rho} d\rho, \tag{11.32}
\]

where \( q = k(\hat{s}_p - \hat{s}_o) \) and \( X_o(\rho - \rho_o) \), \( X_p(\rho - \rho_p) \) are the beam profiles within \( \Pi \), centered at \( \rho_o \) and \( \rho_p \), as indicated in Fig. 11.4. In a MC simulation we use a simplified version of the beam selection. We include the polarization transfer in the rigorous way, but simplify the beam overlap. Each of the pixels of the CCD array in Fig. 11.4 corresponds to a rectangular pixel on the \( \Pi \)-plane, whose area \( \Delta x_o \Delta y_o \) corresponds to the effective cross-section \( \pi a_o^2 \) of the observation beam. The emerging photon states are beams with some effective cross-section \( \pi(\delta d)^2 \), which we can choose more or less at will, since \( \delta d \) is not used in MC simulation. Therefore we do not bother with the calculation of the overlap integral. A photon is accepted if two conditions are met: i) in its ultimate jump the photon crosses the interface within the pixel area \( \Delta x_o \Delta y_o \) around \( x_o, y_o \). ii) the angle \( \theta \) between the photon propagation vector \( \hat{s}_p \) and the receiver axis \( \hat{s}_o \) is lower than the angular aperture of the receiver. Now we insert polarization filters and measure the polarization, but this we discussed extensively in section 4 and 6.

### 11.7 From path probabilities to Monte Carlo

We summarize the formal development of the path model. The time-averaged probability \( \bar{P}_d(\Pi) \) that a photon which was injected in the state \( |\Pi\rangle \) will be detected in a single-mode detection channel can be expressed as \( \bar{P}_d = \langle |E'_x|^2 \rangle_t + \langle |E'_y|^2 \rangle_t \). Here \( E'_x \) and \( E'_y \) are the components of the detected Jones vector, which may be regarded as a vectorial detection amplitude. On its way from the laser beam through the tissue into the receiver the photon may propagate on many different paths \( p \), each of them consisting of different numbers \( np \) of scattering events at different positions \( \{r_i\}_p \) in the sample. The total detection amplitude is therefore the sum of the amplitudes of individual paths. In our case this so-called path integral ([79]) reads:

\[
\left( \begin{array}{c} E'_x \\ E'_y \end{array} \right) = F \sum_p^{N_p} T_{po} T_{pt} \left\{ \Pi^{np}_{jp} \bar{P}_d \right\} \left( \begin{array}{c} E'_l \\ E'_l \end{array} \right) \tag{11.33}
\]
Here $c_x^l$ and $c_y^l$ are the components of the input state $|l\rangle$ such that $\langle l|l \rangle = 1$. The product $\prod_{j=1}^{np} F_{jp}$ represents a chain of $np$ scattering events with segment amplitudes $F_{jp}$. Since we are interested in time averages, and since we employ the decorrelation approximation, we use the pre-averaged segment amplitudes from Sec. 11.5. The matrix $T_{pt}$ represents a special scattering event, namely the transmission of the photon across the tissue-air interface (recall Sec. 3.5.1). The matrix $T_{po}$ is the beam filter matrix, Eq. 11.31. Finally, $F$ is one of the filters from Eqs. 3.16, 3.17 and 3.18.

The key task in Feynman’s path analysis would be to calculate the path integral. In our case this would first mean to find all possible paths that lead from the laser beam into the detector. This is a tremendous work, which is only slightly simplified by the fact that we do not need to add the amplitudes but only the probabilities. (See [79] for an attempt to apply the path integral approach to radiative transfer.) The problem is, that when searching for all paths, we would find a great number of paths which are highly improbable. When following such a path, a photon makes it into the detector with negligible probability. Here is where Monte Carlo comes in. Instead of searching for all paths, we select a relatively small sample of paths that are probable enough to bring some of the photons into the detector. Thus, instead of searching for all paths and then weighting them with their probabilities, we sample the paths by choosing the distances $d_i$, the absorption events, and the angles $\phi_i, \theta_i$ randomly, but according to their probability distributions. Keep also in mind that in each step of the path one must examine if the photon makes an attempt to cross the tissue-air interface. If yes, a special sampling procedure is required to decide if the photon is reflected from the interface or if it is transmitted.

Examining Eq. 11.27 one finds a feature that significantly distinguishes the polarized MC simulation from the scalar version and which makes sampling of the angles $\theta$ and $\phi$ in polarized MC technically more demanding: the polarized phase function $p(\theta, \phi; \pi_{i-1})$ depends on the polarization state of the incoming photon, which keeps changing in every segment of the path. Moreover, different initial polarization states $|l\rangle$ prefer different paths. This feature has two practical consequences: First, the only practicable way to sample the angles $\theta$ and $\phi$ is the acceptance-rejection method [80]. However, the problem is to find an efficient instrumental distribution, i.e. a simple phase function $p'(\theta)$ which forms a tight envelope of $p(\theta, \phi; \pi_{i-1})$ for any incoming state $\pi_{i-1}$. This is quite a challenge if one works with Mie spheres [68], but for our dipole scattering matrix the envelope function can be rather easily deduced. The second consequence is that one needs some patience when simulating the Müller matrix. A new simulation is needed for each of the four required initial polarization states.

Note that each sampling act corresponds to a measurement in the sense of quantum mechanics. Each sampling act makes the probability to reality. The photon takes the new path segment, just as a billiard ball would do. Thus, after each step $i$ of the path (including transition across tissue-air interface) has been sampled, one must renormalize the photon state $|\pi_{i}\rangle$, so that $\langle \pi_{i}|\pi_{i} \rangle = 1$. No sampling is needed in the final detection steps: If the photon makes it into the receiver, then we assign it the weight $\langle \pi_{o}|\Pi F^\dagger F\Pi|\pi_{o} \rangle$, where $F$ is the filter matrix and $\Pi$ the polarization matrix from Eq. 11.31. Note that except for this final step, we refrain from weighting or re-weighting a previously sampled path. In particular, we sample the absorption events, thereby terminating the path, instead of propagating the photon all the way through and then multiplying with $\exp(-\mu_a L)$. We also refrain from re-using a path or its parts. (In particular we generate a new path for each initial state.) Such path re-weighting and re-using is more likely to cause statistical artifacts, rather than a true improvement of the accuracy.
Bibliography


