2 Crystal Optics

2.1 Polarization

In this section, we review some basic properties of electromagnetism in crystalline media. In particular, we introduce the notion and classification of the polarization of an EM-wave in a homogeneous medium. We consider monochromatic plane waves (timeharmonic plane waves) that propagate in z-direction.

$$\mathbf{E}(z,t) = \operatorname{Re}\left\{\mathbf{A}e^{\mathrm{i}(\mathbf{k}\mathbf{r}-\omega t)}\right\}$$
(2.1)

with complex amplitude
$$\mathbf{A} = A_x \hat{\mathbf{x}} + A_y \hat{\mathbf{y}}, \quad A_x, A_y \in \mathbb{C}.$$
 (2.2)

In the above equation, we have introduced the unit vectors a long the x- and y-direction, $\hat{\mathbf{x}}$ and $\hat{\mathbf{y}}$, respectively. Furthermore, we decompose the wave vector \mathbf{k} into its magnitude k and direction $\hat{\mathbf{k}}$. Without loss of generality, we assume that the wave vector points along the z-axis (we can always choose the orientation of our coordinate system in that way):

$$\mathbf{k} = k \cdot \hat{\mathbf{k}}$$
 and $\hat{\mathbf{k}} = \begin{pmatrix} 0\\0\\1 \end{pmatrix} = \hat{\mathbf{z}}$ (2.3)

also
$$A_i = \underbrace{a_i}_{\in \mathbb{R}} e^{i\varphi_i}$$
, for $i = x, y$. (2.4)

$$\Rightarrow E_x = a_x \cos\left(kz - \omega t + \varphi_x\right) \tag{2.5}$$

$$E_y = a_y \cos\left(kz - \omega t + \varphi_y\right) \tag{2.6}$$

$$\Rightarrow \frac{E_x^2}{a_x^2} + \frac{E_y^2}{a_y^2} - 2\cos\varphi \frac{E_x E_y}{a_x a_y} = \sin^2\varphi \tag{2.7}$$

with
$$\varphi = \varphi_x - \varphi_y.$$
 (2.8)

The above manipulations give the following picture: If we keep z fixed, then the direction of **E** rotates (as a function of time t) in the xy-plane and traces out the so-called polarization ellipse. There are two special cases, which we discuss below:

Linearly polarized light :

$$a_x = 0 \quad \text{or} \quad a_y = 0$$

or
$$\varphi = 0 \quad \text{or} \quad \varphi = \pi$$

(2.9)

$$\Rightarrow E_y = \pm \left(\frac{a_y}{a_x}\right) E_x \tag{2.10}$$

In this case, the polarization ellipse degenerates into a line.

Circularly polarized light :

$$\varphi = \pm \frac{\pi}{2}, \quad a_x = a_y = a_0 \tag{2.11}$$

$$\Rightarrow E_x^2 + E_y^2 = a_0^2 \tag{2.12}$$

In this case, the polarization ellipse becomes a circle and can be traced out either in a clockwise or counterclockwise fashion. These are the so-called Right- and Left-circularly polarized light waves.

In all cases, there are (for a given wave vector) two linearly independent components and one can combine this into a matrix description which is the so-called Jones matrix representation that works with (2×2) matrices. This is somewhat analogous to the description of electron spins in Quantum Mechanics.

2.2 Optics of (simple) anisotropic media

In order to understand the principal behavior of anisotropic materials, it is best to first ignore any dispersive property (frequency dependence) of the material. Owing to the Kramers-Kronig relation the material will then be lossless. This is what we mean by *a simple anisotropic material*.

In mathematical terms, the meaning of 'simple' is $\mathbf{D} = \epsilon_0 \epsilon \mathbf{E}$, $\mathbf{B} = \mu_0 \mathbf{H}$. Here, we have introduced the dielectric tensor $\epsilon_{=}$, i.e., a (3 × 3) matrix that has the following properties :

- Non-dispersiveness implies lossless medium (by Kramers-Kronig relation), i.e., $\epsilon_{=}$ is Hermitian.
- For now, we assume that ϵ_{ij} is real symmetric, i.e $\epsilon_{ij} = \epsilon_{ji}$. This assumption will be lifted in Sec. 2.3.2.

- Because plane wave propagation is the only allowed form of propagation, ϵ is real symmetric and positive definite.
- The above property implies that ϵ has 3 positive eigenvalues $\epsilon_1, \epsilon_2, \epsilon_3$ and corresponding eigenvectors (principal axes) that are mutually orthogonal.

For a given material, the values $\epsilon_1, \epsilon_2, \epsilon_3$ and the corresponding principal axes are determined by the microscopic arrangements and symmetries of the material's constituents (atoms, molecules, etc). It is one of the tasks of solid-state physics to determine these quantities. In optics we take these quantities as given input.

We consider time-harmonic plane waves, i.e., **E**, **D**, **B**, **H** depend on **r** and t through the common factor $e^{i(k\hat{\mathbf{k}}\cdot\hat{\mathbf{r}}-\omega t)}$ as in $\mathbf{E}(\mathbf{r},t) = \mathbf{E}_0 e^{i(k\hat{\mathbf{k}}\cdot\hat{\mathbf{r}}-\omega t)}$ with a constant vector \mathbf{E}_0 (same goes for **D**, **B**, and **H**). We also ignore (free) charges and (free) currents, hence we set $\rho = 0$ and $\mathbf{j} = 0$. Then we can deduce from the Maxwell equations (1.45) and (1.47):

$$k\mathbf{k} \times \mathbf{H} + \omega \mathbf{D} = 0 \Rightarrow \mathbf{D} \perp \mathbf{k}, \mathbf{H}$$
 (2.13)

$$k\hat{\mathbf{k}} \times \mathbf{E} - \mu_0 \omega \mathbf{H} = 0 \Rightarrow \mathbf{H} \perp \hat{\mathbf{k}}, \mathbf{E}$$
 (2.14)

$$\Rightarrow \mathbf{H} \perp \underbrace{\hat{\mathbf{k}}, \mathbf{E}, \mathbf{D}}_{\text{coplanar}}$$
(2.15)

$$\hat{\mathbf{k}} \times \mathbf{H} = \frac{k}{\mu_0 \omega} \underbrace{\hat{\mathbf{k}} \times \left(\hat{\mathbf{k}} \times \mathbf{E} \right)}_{= -\mathbf{E} + \left(\mathbf{E} \cdot \hat{\mathbf{k}} \right) \hat{\mathbf{k}}} = -\frac{\omega}{k} \mathbf{D}$$
(2.16)

$$\Rightarrow \mathbf{D} = \frac{k^2}{\mu_0 \omega^2} \left[\mathbf{E} - \left(\mathbf{E} \cdot \hat{\mathbf{k}} \right) \hat{\mathbf{k}} \right]$$
(2.17)

Eq. (2.17) implies $\hat{\mathbf{k}} \cdot \mathbf{D} = 0$, i.e., the divergence condition (2.13) for the displacement field is consistent with (2.17).

Furthermore, $\mathbf{E} \cdot \mathbf{k}$ can be non-zero now. If the dispersion relation is known, this allows to determine the polarization (i.e., the direction) of \mathbf{D} .

Note:

$$\mathbf{H} = \frac{k}{\mu_0 \omega} \hat{\mathbf{k}} \times \mathbf{E} \tag{2.18}$$

$$\Rightarrow \mathbf{S} = \frac{1}{2} \operatorname{Re}(\mathbf{E} \times \mathbf{H}^*)$$
 (2.19)

$$= \frac{1}{2} \frac{k}{\mu_0 \omega} \operatorname{Re}\left[\left(\mathbf{E} \cdot \mathbf{E}^* \right) \hat{\mathbf{k}} - \left(\mathbf{E} \cdot \hat{\mathbf{k}} \right) \mathbf{E}^* \right]$$
(2.20)

 \Rightarrow The Poynting-vector **S** is not necessarily parallel to the wave vector **k**.

We now insert the constitutive relation $\mathbf{D} = \epsilon_0 \epsilon \mathbf{E}$ into Eq. (2.17). This gives

$$\underbrace{\left(\underbrace{\epsilon}_{=} - \frac{c^{2}k^{2}}{\omega^{2}}\mathbb{1} + \frac{c^{2}k^{2}}{\omega^{2}}\hat{\mathbf{k}} : \hat{\mathbf{k}}\right)}_{\text{det}=0 \text{ for non trivial solutions}} \mathbf{E} = 0 \text{ independent of Coordinate system}$$
(2.21)

Without loss of generality, we can assume that we are in a coordinate system that coincides with the material's principal axes:

$$\begin{aligned} \epsilon &= \begin{pmatrix} \epsilon_1 & \\ & \epsilon_2 & \\ & & \epsilon_3 \end{pmatrix}, \qquad \hat{k} = \begin{pmatrix} k_1 \\ k_2 \\ k_3 \end{pmatrix}. \end{aligned}$$
 (2.22)

Then, the requirement of the determinant being zero gives the celebrated **Fresnel Equa**tion:

$$\left(\frac{c^2k^2}{\omega^2}\right)\left(\frac{\epsilon_1c^2k_1^2}{\omega^2} + \frac{\epsilon_2c^2k_2^2}{\omega^2} + \frac{\epsilon_3c^2k_3^2}{\omega^2}\right)$$
(2.23)

$$-\left(\frac{c^2k_1^2}{\omega^2}\epsilon_1\left(\epsilon_2+\epsilon_3\right)+\frac{c^2k_2^2}{\omega^2}\epsilon_2\left(\epsilon_1+\epsilon_3\right)+\frac{c^2k_3^2}{\omega^2}\epsilon_3\left(\epsilon_1+\epsilon_2\right)\right)$$
(2.24)

$$+ \epsilon_1 \epsilon_2 \epsilon_3 = 0 \qquad (2.25)$$

where, $k^2 = k_1^2 + k_2^2 + k_3^2$.

2.3 Discussion of general properties

We proceed to discuss the properties of the Fresnel equations by first considering its general properties and then consider successively more difficult special cases.

For given $\epsilon_1, \epsilon_2, \epsilon_3$:

- For fixed ω and given direction $\hat{\mathbf{k}}$ there generally exist two distinct values for the wave number $k \in \mathbb{R}$ that solve the Fresnel equation. Exceptions occur at so-called *double points*.
- For fixed k and given direction of $\hat{\mathbf{k}}$ there exist two frequencies ω that solve the Fresnel equation.

One can define a refractive index vector by $\mathbf{n} := c \frac{k}{\omega} \hat{\mathbf{k}}$ which is also called the *slowness-vector* for which sometimes also the notation \mathbf{s} is used.

One can define a refractive index $n := |\mathbf{n}|$ for the direction $\hat{\mathbf{k}} = \frac{ck}{\omega}\mathbf{n}$.

Note:

- The index of refraction is a wave property, not a material property.
- $\epsilon_1, \epsilon_2, \epsilon_3$ and the principal axes are material properties and not wave properties.
- The dispersion relation $\omega = \omega(\mathbf{k})$ (i.e., the solution of the Fresnel equation) characterizes the wave completely, so that there is no real need to introduce a refractive index.

However, if you do use the notion of a refractive index, be aware that a refractive index is a *wave* and not a *material* property.

Note:

• Strictly speaking, the refractive index introduced above corresponds to the phase of the wave, i.e., (for given $\hat{\mathbf{k}}$) it relates the phase velocity $\frac{\omega}{k}$ to the vacuum speed of light.

Therefore, the correct statement should be that $n = \frac{ck}{\omega}$ is a phase index, denoted as $n_{\rm ph}$.

In general, we have in anisotropic media that

$$\mathbf{v}_{\text{group}} \not\parallel \mathbf{v}_{\text{phase}}$$
 (2.26)

and

$$|\mathbf{v}_{\text{group}}| \neq |\mathbf{v}_{\text{phase}}|$$
 (2.27)

Therefore, one could equally well define a group index $n_{\rm gr} := \frac{c}{\left|\frac{\partial \omega}{\partial \mathbf{k}}\right|}$, but this would —

in general — be for another direction !

In terms of the dispersion relation (recall that we are in lossless non-magnetic media), we have less confusion with

$$\mathbf{v}_{\text{phase}} = \hat{\mathbf{k}} \frac{\omega\left(\mathbf{k}\right)}{k} \tag{2.28}$$

and

$$\mathbf{v}_{\text{group}} = \frac{\partial \omega(\mathbf{k})}{\partial \mathbf{k}} = \frac{\langle \mathbf{S} \rangle}{\langle w \rangle} = \mathbf{v}_{\text{E}}, \qquad (2.29)$$

where the vector of the energy transport velocity is sometimes $\mathbf{v}_{\rm E}$ called the *ray vector*.

The above property that the group velocity is identical to the energy transport velocity (and thus parallel to the Poynting vector) only applies to lossless and nonmagnetic systems. However, if these criteria are satisfied for all constituent materials, the statement also applies to periodically structured materials such as Bragg stacks or Photonic Crystals:

$$\mathbf{v}_{\text{group}} \parallel \mathbf{S} = \frac{1}{2} \text{Re}(\mathbf{E} \times \mathbf{H}^*)$$
(2.30)

??writeme: check reference The proof is found in the paper by P. Yeh JOSA **69**, 742 (1979) (see Fig. 2.13).

writeme

iteme

iteme

Note: Remember the assumptions: ??writeme: check this table!

lossless	non-magnetic
ϵ is real symmetric	$\mu = 1$
and positive definite	

If any of the two does not apply, some or all of the above considerations do not hold.

Isotropic material (Cubic symmetry): For

$$\epsilon_1 = \epsilon_2 = \epsilon_3 = \epsilon \tag{2.31}$$

the Fresnel Equation becomes

$$\epsilon \left(\frac{c^2 k^2}{\omega^2} - \epsilon\right)^2 = 0. \tag{2.32}$$

Dispersion relation has shape of two identical spheres, one for each transverse polarization:**??writeme:** check next two lines

$$\mathbf{v}_{\rm ph} = \mathbf{v}_{\rm gr} = \frac{c}{\sqrt{\epsilon}} \quad \text{where} \quad \sqrt{\epsilon} = n_{\rm ph} = n_{\rm gr},$$
 (2.33)

i.e. $\mathbf{n} \parallel \mathbf{S}$.

Optically uni-axial materials (Trigonal, tetragonal, hexagonal symmetry): Here we have

$$\epsilon_1 = \epsilon_2 = \epsilon_\perp. \tag{2.34}$$

 $\epsilon_3 > \epsilon_{\perp}$: Positive uni-axial

 $\epsilon_3 < \epsilon_{\perp}$: Negative uni-axial

The Fresnel Equation becomes

$$\left(\frac{c^2k^2}{\omega^2} - \epsilon_{\perp}\right) \left(\epsilon_{\perp} \frac{c^2k_{\perp}^2}{\omega^2} + \epsilon_3 \frac{c^2k_3^2}{\omega^2} - \epsilon_{\perp}\epsilon_3\right) = 0.$$
(2.35)

In the above equation the solutions take on following shapes:

$$\left(\frac{c^2k^2}{\omega^2} - \epsilon_{\perp}\right) \Rightarrow \text{ one sphere } \Rightarrow \text{ ordinary wave}$$
(2.36)

$$\left(\epsilon_{\perp} \frac{c^2 k_{\perp}^2}{\omega^2} + \epsilon_3 \frac{c^2 k_3^2}{\omega^2} - \epsilon_{\perp} \epsilon_3\right) \Rightarrow \quad \text{one ellipsoid} \quad \Rightarrow \quad extraordinary \text{ wave} \qquad (2.37)$$

The ellipsoids may be

positive uni-axial: circumscribed

negative uni-axial: inscribed

See Figs. 2.1, 2.2 and 2.3.





Special cases:

- k || optical axis: Same phase and group velocity, same direction of k, S
 For linearly polarized waves, the orientation of the polarization remains unchanged.
- $\hat{\mathbf{k}} \perp$ optical axis: Same direction of $\hat{\mathbf{k}}$, \mathbf{S} , different magnitudes for phase and group velocity. An initially linearly polarized wave will become elliptically polarized.

General case:

• $\hat{\mathbf{k}} \not\mid \mathbf{S}$: Different magnitudes of phase and group velocity. Two rays.

Optically bi-axial materials (Orthorhombic, monoclinic, triclinic symmetry): We order the values as $\epsilon_1 < \epsilon_2 < \epsilon_3$ and have to consider the full Fresnel equation

$$\left(\frac{c^2k^2}{\omega^2}\right)\left(\frac{\epsilon_1c^2k_1^2}{\omega^2} + \frac{\epsilon_2c^2k_2^2}{\omega^2} + \frac{\epsilon_3c^2k_3^2}{\omega^2}\right) - \left(\frac{c^2k_1^2}{\omega^2}\epsilon_1\left(\epsilon_2 + \epsilon_3\right) + \frac{c^2k_2^2}{\omega^2}\epsilon_2\left(\epsilon_1 + \epsilon_3\right) + \frac{c^2k_3^2}{\omega^2}\epsilon_3\left(\epsilon_1 + \epsilon_2\right)\right) + \epsilon_1\epsilon_2\epsilon_3 = 0 \quad (2.38)$$

To gain some insight, we first set $k_3 = 0$. Then

$$\left(\frac{c^2k^2}{\omega^2} - \epsilon_3\right) \left(\epsilon_1 \frac{c^2k_1^2}{\omega^2} + \epsilon_2 \frac{c^2k_2^2}{\omega^2} - \epsilon_1\epsilon_2\right) = 0.$$
(2.39)

In the above equation we have

$$\left(\frac{c^2k^2}{\omega^2} - \epsilon_3\right) \Rightarrow \text{Circle (in } k_1, k_2)$$
$$\left(\epsilon_1 \frac{c^2k_1^2}{\omega^2} + \epsilon_2 \frac{c^2k_2^2}{\omega^2} - \epsilon_1\epsilon_2\right) \Rightarrow \text{Ellipse (in } k_1, k_2)$$



Figure 2.2: Intersection of the \mathbf{k} surface with the yz-plane for a uni-axial crystal.

Owing to the ordering of $\epsilon_1, \epsilon_2, \epsilon_3$, the ellipse lies inside circle.

We next consider the analogous cases

- $k_1 = 0$: We obtain the same, except that the circle lies inside ellipse,
- $k_2 = 0$: We obtain that the circle and ellipse intersect at two opposite pairs of points (see Fig. 2.4).

Those complicated dispersion surfaces have been considered by Sir Rowan Hamilton shortly before the Maxwell equations were introduced. Sir Hamilton suggested that this should lead to the phenomenon of inner and outer conical refraction, an effect which subsequently has been confirmed by experiment (See Fig. 2.5) **??writeme: explain annotations in plot fig:ConicalRefraction**

iteme

2.3.1 Reflection and Refraction at Interfaces

The general setup is shown in Fig. 2.6. We have materials with different dispersion relations and consider the interface conditions:

Kinematics:

- Frequency conserved.
- Wave vector parallel to surface conserved.

Dynamics:

- reflected and transmitted waves have to take the incoming energy <u>away</u> from the surface (valid for any dispersion relation).



Figure 2.3: (a) Ordinary wave, polarized perpendicular to the k-z plane. (b) Extraordinary wave with a polarization vector in the k-z plane.

Note:

• No statement regarding the normal component of the wave vector is made.

In the absence of surface charges and surface currents it follows that

$$\underbrace{ \begin{array}{ccc} \hat{\mathbf{n}} & \times & (\mathbf{E_1} - \mathbf{E_2}) & = 0 \\ \hat{\mathbf{n}} & \times & (\mathbf{H_1} - \mathbf{H_2}) & = 0 \end{array} }_{\text{Tangential components of } \mathbf{E}, \mathbf{H} \text{ continuous}} \Rightarrow \underbrace{ \begin{cases} \hat{\mathbf{n}} & \cdot & (\mathbf{D_1} - \mathbf{D_2}) & = 0 \\ \hat{\mathbf{n}} & \cdot & (\mathbf{B_1} - \mathbf{B_2}) & = 0 \\ \text{Normal components of } \mathbf{B}, \mathbf{D} \text{ continuous} \end{cases} }$$

-Implies normal component of $\mathbf{S} = \mathbf{E} \times \mathbf{H}$ is continuous across the interface.

-Otherwise energy would be accumulated.

Based on the above discussion, we have to make the following ansatz for the media 1 and 2.

Ansatz medium 1:

$$\mathbf{E} = \hat{\mathbf{e}} e^{i(\mathbf{k}\mathbf{r}-\omega t)} + \sum_{\beta=1}^{2} A_{\beta} \hat{\mathbf{e}}_{\beta} e^{i\left(\mathbf{k}_{\beta}'\mathbf{r}-\omega t\right)}$$
(2.40)

Ansatz medium 2:

$$\mathbf{E} = \sum_{\gamma=1}^{2} B_{\gamma} \hat{\mathbf{e}}_{\gamma} \mathrm{e}^{\mathrm{i} \left(\mathbf{k}_{\gamma}^{\prime\prime} \mathbf{r} - \omega t \right)}$$
(2.41)

The coefficients in these ansatzes have to be determined by, for instance, equating the tangential components of the electric and magnetic field. See Figs. 2.7, 2.8 and 2.9.



Figure 2.4: Wave surfaces for optically bi-axial crystals.

2.3.2 A potpourri of optically anisotropic materials

Liquid Crystals

Liquid Crystals consist of cigar-shaped molecules (rotational ellipsoids) that have different dipole moments (or polarizabilities) along the various symmetry directions. For sufficiently low temperatures these molecules order automatically in various types of structures. Most common are nematic, smectic, cholesteric liquid crystals (see Fig. 2.15). For instance, for a nematic liquid crystal the molecules point — on average — along a certain direction. More generally, liquid crystals may be characterized by a local director field $\hat{\mathbf{n}}(\mathbf{r})$ that describes the average orientation of the molecules. The different dipole moments lead — upon averaging — to a dielectric tensor of the form

$$\epsilon = \epsilon_{\perp} 1 + \left(\epsilon_{\parallel} - \epsilon_{\perp}\right) \mathbf{\hat{n}} : \mathbf{\hat{n}}.$$
(2.42)

Optical Activity

A medium characterized by (harmonic time dependence assumed)

$$\mathbf{D} = \epsilon_0 \epsilon \mathbf{E} + \mathrm{i}\omega \epsilon_0 \xi \mathbf{B} \tag{2.43}$$

$$=\epsilon_0 \epsilon \mathbf{E} - \epsilon_0 \xi \nabla \times \mathbf{E} \tag{2.44}$$

is called optically active.

Such a relation arises in molecular structures with a helical character. In these structures, a time-varying magnetic flux density **B** induces a circulating current that sets up an electric dipole moment (and hence polarization) proportional to $i\omega \mathbf{B} = -\nabla \times \mathbf{E}$. The optically active medium is spatially dispersive, meaning that the relation between $\mathbf{D}(\mathbf{r})$ and $\mathbf{E}(\mathbf{r})$ is non-local in space.

For a plane wave $\mathbf{E}(\mathbf{r}) = \mathbf{E}_0 e^{i\mathbf{k}\mathbf{r}}$ we have $\mathbf{D}(\mathbf{r}) = \mathbf{D}_0 e^{i\mathbf{k}\mathbf{r}}$ so that,

$$\mathbf{D}_0 = \epsilon_0 \epsilon \mathbf{E}_0 - \mathrm{i} \epsilon_0 \mathbf{G} \times \mathbf{E}_0 \tag{2.45}$$



Figure 2.5: (a) Principle of inner conical refraction. (b) Principle of outer conical refraction.



Figure 2.6: General setup for reflection at an interface.

Here, we have introduced the gyration vector $\mathbf{G} = \xi \mathbf{k}$.

Note: As can be seen here, \mathbf{G} changes sign under reversal of the propagation direction, i.e., change of the direction of the wave vector.

We can now express this as

$$\mathbf{D}_0 = \epsilon_0 \underline{\epsilon} \mathbf{E}_0 \tag{2.46}$$

where

$$\begin{aligned}
\epsilon &= \begin{pmatrix} \epsilon & \mathrm{i}G_z & -\mathrm{i}G_y \\ -\mathrm{i}G_z & \epsilon & \mathrm{i}G_x \\ \mathrm{i}G_y & -\mathrm{i}G_x & \epsilon \end{pmatrix}.
\end{aligned}$$
(2.47)



Figure 2.7: Determination of the angles of refraction by matching projections of the \mathbf{k} vectors in air and in a uni-axial crystal.



Figure 2.8: Double refraction at normal incidence.



Figure 2.9: Double refraction through an anisotropic plate. The plate serves as a polarizing beam splitter.



Figure 2.10: An index ellipsoid where the principal axes are along the (x,y,z)-axes and the principal refractive indices are given by n_1, n_2, n_3 .



Figure 2.11: Normal modes determined from the index ellipsoid.



Figure 2.12: An octant of the k surface for (a) a bi-axial crystal, (b) a uni-axial crystal and (c) an isotropic crystal.



Figure 2.13: Rays and wave fronts for (a) a spherical ${\bf k}$ surface and (b) a non-spherical ${\bf k}$ surface.



Figure 2.14: (a)Variation of the refractive index $n(\theta)$, (b) direction of the **D** and **E** vectors for the ordinary and extraordinary wave.



Figure 2.15: Molecular orientation of various types of liquid crystals: (a) nematic, (b) smectic and (c) cholesteric.

In particular, for circularly polarized light that propagates along the z-axis with

$$\mathbf{k} = \begin{pmatrix} 0\\0\\k \end{pmatrix} \text{ and } \mathbf{E}_0 = \begin{pmatrix} E_0\\\pm iE_0\\0 \end{pmatrix}, \qquad (2.48)$$

we obtain

$$\mathbf{D}_0 = \begin{pmatrix} D_0 \\ \pm \mathbf{i} D_0 \\ 0 \end{pmatrix} \tag{2.49}$$

where

$$\mathbf{D}_{0} = \epsilon_{0} \underbrace{(\epsilon \pm \xi k)}_{=n_{\pm}^{2}} \mathbf{E}_{0} \tag{2.50}$$

In other words, we obtain different phase velocities for left- and right-circularly polarized light. This phenomenon is called circular birefringence and should not be confused with the case of linear birefringence. However, similar to linear birefringence, this leads to rotation of plane of polarization for linear polarized light upon propagation. Note that when the direction of propagation is reversed, the direction of polarization rotation is also reversed.

Faraday Effect

Certain materials show for plane waves and for propagation in a given direction a similar polarization rotation when placed in a <u>static</u> magnetic field \mathbf{B}_{stat} .

$$\mathbf{D} = \epsilon_0 \epsilon \mathbf{E} + \mathbf{i} \omega \epsilon_0 \gamma \mathbf{B}_{\text{stat}} \times \mathbf{E}.$$
 (2.51)

Here, γ is called the magneto-gyration coefficient.

The above may be rewritten as,

$$\mathbf{D} = \epsilon_0 \epsilon \mathbf{E} + \mathbf{i} \omega \underbrace{\mathbf{G}_{\text{Faraday}}}_{\gamma \mathbf{B}_{\text{stat}}} \times \mathbf{E} = \epsilon_0 \epsilon \mathbf{E}_{=} \tag{2.52}$$

$$\epsilon = \begin{pmatrix} \epsilon & i\omega G_z & -i\omega G_y \\ -i\omega G_z & \epsilon & i\omega G_x \\ i\omega G_y & -i\omega G_x & \epsilon \end{pmatrix}$$
 (2.53)

Note:

- 1. Despite the apparent similarity to the case of optical activity, Faraday-Activity is quite different. For instance, upon change of propagation direction, G_{Faraday} does not change sign!
 - a) The polarization rotation is the same direction for forward *and* backward propagating modes. Therefore, Faraday-active materials break time-reversal symmetry!
 - b) A Faraday medium is spatially non-dispersive.
- 2. Sometimes the Faraday-effect is described in terms of the Verdet constant V, given as

$$V \approx -\frac{\pi\gamma}{\lambda_0 n_0} \tag{2.54}$$

where
$$n_0 = \sqrt{\epsilon}$$
 and $\lambda_0 = 2\pi \frac{c}{\omega}$ (2.55)

Typical Faraday-active materials include

YIG (Yttrium-Iron-Garnet)

TGG (Terbium-Gallium-Garnet)

TbAIG (Terbium-Aluminum-Garnet)

In general, Faraday-active materials that contain rare-earth elements show the largest effects, since the f-electrons of the rare-earth elements react best to magnetic fields.

For example:

$$V = -1.16 \frac{\min}{\text{cmOe}} \quad \text{at } \lambda_0 = 500 \text{nm}$$
 (2.56)

Recall that in vacuum 1 Oe = 1 Gauss = 10^{-4} T.

Device Application:

- Optical Isolators (the optical analogon to an electrical diode, see Fig. 2.16)
- Circulators (see, e.g., Wikipedia)



Figure 2.16: Realization of an optical insulator by using a Faraday rotator. Light is transmitted in one direction (a), while the propagation in the other direction is blocked (b).

3 Diffraction Theory

This chapter is based largely on Goodman, Introduction to Fourier Optics [3] and Römer, Theoretical Optics [4]. This chapter deals with monochromatic (only one single frequency involved) scalar diffraction theory exclusively. Therefore, only the scalar complex valued wave envelope $U(\mathbf{r})$ in the stationary case is considered (i.e., U varies in space but not in time), which can represent any electric or magnetic field component in the following, since all these components obey the same differential equation here.

3.1 Phenomenology: The Principles of Huygens and Fresnel

The simplest theory regarding light propagation is geometrical optics, where light is treated as rays which propagate in straight lines in vacuum. However, this is just a crude approximation and neglects the rich physical behavior due to the wave nature of light. This wave character of light is responsible for effects like diffraction and interference phenomena. Understanding diffraction in the wave picture is the subject of classical diffraction theory.

According to Huygens' principle, each point of the surface of a wave front (the surface of constant phase F_0) may be considered as the origin of a new elementary (spherical) wave. New surfaces of constant phase at a later instant of time are obtained as the enveloping surface of the family of phase surfaces of these elementary waves (see Fig. 3.1). This enables one to construct the full wave pattern in all of space, if only one single surface (of constant phase) of the wave is initially known.

However, this picture immediately raises the question: Why does light propagate along lines in homogeneous materials then, which is properly described by the obviously inferior geometrical optics approach? In fact, this puzzle has lead Newton to abandon a wave theory for light in favor of a particle theory. The problem has been solved by Augustin Jean Fresnel in 1818 (before Maxwell's equations!): Consider a monochromatic point source in Fig. 3.2, located at some point \mathbf{x}_0 in space and emitting a wave with wavelength λ and corresponding wavenumber $k = \frac{2\pi}{\lambda}$. We want to determine what is observed at the point \mathbf{x} ? Therefore, imagine spherical shells around \mathbf{x} with radii $r_n = n\lambda + \rho$. This decomposes the spherical wave emitted by the point source into so-called Fresnel zones Z_n .

All points in a given zone have the same distance $\rho + (n-1)\lambda \leq r_n(\rho) \leq \rho + n\lambda$ from **x** and can be considered as the origin of elementary waves of the form



Figure 3.1: Elementary waves originating at each point of a non-trivial surface defined by the phase value F_0 of a wave. As these elementary waves propagate, a new non-trivial wave front forms as the envelope defining the surface of equal phase from the superposition of elementary waves. Thus the whole wave front (of complicated shape) propagates in space in terms of simple elementary waves.



Figure 3.2: Construction of Fresnel zones

$$C\kappa(\chi) \frac{\mathrm{e}^{\mathrm{i}k|\mathbf{x}-\mathbf{y}|}}{|\mathbf{x}-\mathbf{y}|} \qquad \mathbf{y} \in Z_n,$$
(3.1)

where the terms mean

- C: some normalization constant,
- $\kappa(\chi)$: Inclination factor that allows that the strength of the elementary wave emitted in **y** may depend on the angle χ between the normal direction of the wave surface and the direction $\mathbf{x} - \mathbf{y}$.

In each zone, the phase changes by 2π and for small wavelengths λ (compared to the distance $|\mathbf{x} - \mathbf{x}_0|$) these zones are small.

Fresnel's Argument:

Elementary waves from the outer half of each zone cancel with the elementary waves from the inner half of the next zone.

- Only the contributions from the inner half of the very first zone and the outer half of the very last zone remain.
- Both regions are small (for small λ) which explains the almost straight propagation of light.

The prerequisite of small wavelengths λ is exactly the justification for the use of geometrical optics. Hence, the theories are compatible and do not contradict each other, as geometrical optics is contained as a special case in wave optics.

However, one problem remains in this argument: We do not observe the backward wave emanating from the last zone. Fresnel's ad-hoc solution was the ansatz

$$\kappa(\chi) = 0 \quad \text{for} \quad \chi \ge \frac{\pi}{2}.$$
(3.2)

If we generalize this to the case of arbitrary wave fronts F_0 of wave fields $U(\mathbf{r})$, we see that the principles of Huygens and Fresnel correspond to an identity of the form

$$U(\mathbf{r}) = C \int_{F_0} \mathrm{d}s' \; \kappa(\chi) U(\mathbf{r}') \frac{\mathrm{e}^{ik|\mathbf{r}-\mathbf{r}'|}}{|\mathbf{r}-\mathbf{r}'|},\tag{3.3}$$

which is a surface integral with χ depending both on **r** and **r**'. The integration range F_0 in the above integral represents the surface of the wavefront. It is the task of diffraction theory to put the above phenomenology on a solid foundation and to derive consequences.

Remark: Fresnel presented his paper to a prize committee of the French Academy of Science. His theory was strongly disputed by Simeon Denis Poisson (member of the committee) who showed, that it predicted the existence of a bright spot at the center of the shadow of an opaque disk (absurd in geometrical optics). Francois Arago (chair of the prize committee) performed such an experiment and found the predicted spot. Fresnel got the prize, and the effect has since been known as 'Poisson's spot'.

3.2 Scalar Diffraction Theory

The basic problem we are going to investigate is this: Given the wave distribution (amplitude and phase) on some surface (e.g., an illuminated aperture or grating), what is the wave distribution at some other point in space (i.e., which interference pattern will we see on a screen behind that aperture/grating)? Therefore, we need to establish a mathematical relationship between the values of a function (obeying a particular differential equation) on a surface of a volume and the values of that function inside the volume.

In the following, we will consider only scalar diffraction theory, i.e. we neglect the vector nature of light. This is an excellent approximation if the coupling between vector components of the EM-field can be ignored. We further assume

• homogeneous material: $\epsilon(\mathbf{r}) \equiv \epsilon$ (constant),

- isotropic material: $\underset{=}{\epsilon} \equiv \epsilon \mathbb{1}$,
- no strong coupling from boundaries.

The last assumption means that we look at the field far away from boundaries and that the influence of these boundaries on the diffraction process is so weak, such that no coupling between the vector components is introduced and thus the scalar theory stays valid.

We will restrict our discussion to monochromatic waves, i.e. the scalar EM-field of our waves is written as

$$U(\mathbf{r},t) = U(\mathbf{r})\mathrm{e}^{-\mathrm{i}\omega t}.$$
(3.4)

This field then obeys the following scalar wave equation

$$\Delta U(\mathbf{r},t) - \frac{\epsilon}{c^2} \partial_t^2 U(\mathbf{r},t) = 0, \qquad (3.5)$$

where $\Delta = \partial_x^2 + \partial_y^2 + \partial_z^2$ is the Laplacian (i.e., the Laplace operator). Evaluating the time derivative by entering the above form of the scalar wave, we get the (scalar) Helmholtz equation

$$(\triangle + k^2)U(\mathbf{r}) = 0 \tag{3.6}$$

with

$$k := \frac{\omega}{c}\sqrt{\epsilon} = \frac{2\pi}{\lambda}.$$
(3.7)

Here, λ denotes the wavelength in the dielectric medium. This is the equation the envelope function $U(\mathbf{r})$ has to obey in the stationary case along with the boundary conditions that are particular to each physical setup.

Now we will derive the connection between function values at different positions in space for this envelope function $U(\mathbf{r})$.

3.2.1 Mathematical Interlude

Green's Theorem

Let $u(\mathbf{r}), g(\mathbf{r})$ be complex valued functions with single-valued and continuous first and second derivatives. V denotes a volume that is bounded by the closed surface S. Then

$$\iiint\limits_{V} \mathrm{d}^{3}r \; (u \triangle g - g \triangle u) = \iint\limits_{S} \mathrm{d}^{2}r \; (u \partial_{n}g - g \partial_{n}u), \tag{3.8}$$

where $\partial_n u$ is the normal derivative in the outward normal direction $\partial_n u := \hat{\mathbf{n}} \cdot \nabla u$ and $\hat{\mathbf{n}}$ is the outward normal direction at each point of the surface S.

Green's Function

Consider the differential equation

$$a_2\partial_x^2 f + a_1\partial_x f + a_0 f = V(x) \tag{3.9}$$

with certain boundary conditions for a given inhomogeneity V(x). The Green's function of the above ODE is defined as the response of the system (whose dynamics are described by the differential equation above) to an impulse $\delta(x - x')$, that is subject to the *same* boundary conditions:

$$a_2 \partial_x^2 G(x, x') + a_1 \partial_x G(x, x') + a_0 G(x, x') = \delta(x - x').$$
(3.10)

Then, for any given right hand side V(x), the solution of the initial problem can simply be obtained by evaluating the integral

$$f(x) = \int dx' G(x, x') V(x').$$
 (3.11)

Note that there are usually various Green's functions G(x, x') that solve the above differential equation, but only one (or a few) that solve the boundary value problem (i.e., solve the differential equation *and* are compatible with the given boundary conditions). Only then does the integral also yield a solution to the boundary value problem for f(x).

In the following, we derive results corresponding to different assumptions about the Green's function of the problem. For the Helmholtz equation in free space (the differential equation) and the field being zero at infinity (the boundary condition), we have

$$G(\mathbf{r}, \mathbf{r}') = \frac{\mathrm{e}^{\mathrm{i}k|\mathbf{r}-\mathbf{r}'|}}{|\mathbf{r}-\mathbf{r}'|}.$$
(3.12)

In optics, Green's function is known under various names:

- Impulse Response
- Point Spread Function
- Green's Function

3.2.2 Integral Theorem of Helmholtz and Kirchhoff

The aim here is a rigorous derivation of the Huygens-Fresnel principle. We use the Green's function defined above, but have to exclude a tiny volume around the point of singularity \mathbf{r} (where G has a pole) since otherwise we could not apply Green's theorem (see Fig. 3.3). Hence, instead of the volume V bounded by the surface S of interest, we look at V' bounded by S', which is basically V without a small sphere of radius ϵ . At the end, however, we will consider the limit of this tiny volume going to zero (i.e., take the limit $\epsilon \to 0$) which does indeed exist and we arrive at an expression for the field valid in all of V.



Figure 3.3: The surfaces S and S_{ϵ} enclosing the volume V' of integration with exemplary outward surface normals $\hat{\mathbf{n}}$.

So, within the volume V', the field U and the Green's function G obey the Helmholtz equation:

$$(\Delta + k^2)G(\mathbf{r}, \mathbf{r}') = 0, \qquad (3.13)$$

$$(\Delta + k^2)U(\mathbf{r}) = 0, \quad \mathbf{r} \in V'.$$
(3.14)

Applying Green's theorem in V' for these two functions yields (Note that the integration as well as the differentiation ∂_n is with respect to the primed spatial coordinates here and in the following!)

$$\iint_{S'} \mathrm{d}s' \left[U(\mathbf{r}')\partial_n G(\mathbf{r}, \mathbf{r}') - G(\mathbf{r}, \mathbf{r}')\partial_n U(\mathbf{r}') \right] = 0, \qquad (3.15)$$

$$\Longrightarrow_{S'=S\cup S_{\epsilon}} - \iint_{S_{\epsilon}} \mathrm{d}s' \left(U\partial_n G - G\partial_n U\right) = \iint_{S} \mathrm{d}s' \left(U\partial_n G - G\partial_n U\right).$$
(3.16)

Here and in the following, we write $ds' = d^2r'$ for the surface integral (this is just notation, no additional math is involved). This has nothing to do in particular with the surface S'.

On S: The normal derivative with respect to \mathbf{r}' of the Green's function is given as

$$G(\mathbf{r}, \mathbf{r}') = \frac{\mathrm{e}^{\mathrm{i}k|\mathbf{r}-\mathbf{r}'|}}{|\mathbf{r}-\mathbf{r}'|},\tag{3.17}$$

$$\partial_n G(\mathbf{r}, \mathbf{r}') = \cos\left(\angle(\hat{\mathbf{n}}, \mathbf{r} - \mathbf{r}')\right) \left(ik - \frac{1}{|\mathbf{r} - \mathbf{r}'|}\right) \frac{e^{ik|\mathbf{r} - \mathbf{r}'|}}{|\mathbf{r} - \mathbf{r}'|}.$$
(3.18)

On S_{ϵ} : Here we can exploit the explicit spherical shape of S_{ϵ} with *constant* radius $\epsilon = |\mathbf{r} - \mathbf{r}'|$ and get

$$G(\mathbf{r}, \mathbf{r}') = \frac{\mathrm{e}^{\mathrm{i}k\epsilon}}{\epsilon},\tag{3.19}$$

$$\partial_n G(\mathbf{r}, \mathbf{r}') \underbrace{=}_{\cos\left(\angle(\hat{\mathbf{n}}, \mathbf{r} - \mathbf{r}')\right) = -1} \frac{\mathrm{e}^{\mathrm{i}k\epsilon}}{\epsilon} \left(\frac{1}{\epsilon} - \mathrm{i}k\right). \tag{3.20}$$

Plugging these relations into (3.16), we find for the left hand side

$$\iint_{S_{\epsilon}} \mathrm{d}s' \left(U\partial_n G - G\partial_n U \right) = \iint_{S_{\epsilon}} \mathrm{d}s' \left[U(\mathbf{r}') \frac{e^{ik\epsilon}}{\epsilon} \left(\frac{1}{\epsilon} - ik \right) - \partial_n U(\mathbf{r}') \frac{e^{ik\epsilon}}{\epsilon} \right]$$
(3.21)

Now, only $U(\mathbf{r}')$ and $\partial_n U(\mathbf{r}')$ still depend on the integration variable \mathbf{r}' , the remaining terms are constant. The sphere with radius ϵ is centered around the position \mathbf{r} . When choosing ϵ smaller and smaller, the integrals over these functions essentially become area of sphere $\times U(\mathbf{r})$ and area of sphere $\times \partial_n U(\mathbf{r})$ by the first mean value theorem for integration. Hence, we get

$$\iint_{S_{\epsilon}} \mathrm{d}s' \left(U\partial_n G - G\partial_n U \right) \stackrel{\epsilon \ll 1}{\simeq} 4\pi\epsilon^2 \left[U(\mathbf{r}) \frac{e^{ik\epsilon}}{\epsilon} \left(\frac{1}{\epsilon} - ik \right) - \partial_n U(\mathbf{r}) \frac{e^{ik\epsilon}}{\epsilon} \right]$$
(3.22)

$$\stackrel{\epsilon \to 0}{\longrightarrow} 4\pi U(\mathbf{r}). \tag{3.23}$$

Thus, after taking the limit $\epsilon \to 0$, (3.16) becomes the following statement valid for all points $\mathbf{r} \in V$:

$$U(\mathbf{r}) = \frac{1}{4\pi} \iint_{S} \mathrm{d}s' \left[\partial_{n} U(\mathbf{r}') \frac{\mathrm{e}^{\mathrm{i}k|\mathbf{r}-\mathbf{r}'|}}{|\mathbf{r}-\mathbf{r}'|} - U(\mathbf{r}') \partial_{n} \frac{\mathrm{e}^{\mathrm{i}k|\mathbf{r}-\mathbf{r}'|}}{|\mathbf{r}-\mathbf{r}'|} \right].$$
(3.24)

Note the remarkable thing, that we have accomplished now. We can compute the function values $U(\mathbf{r})$ for $\mathbf{r} \in V$ when all we know are the function values $U(\mathbf{r}')$ on the boundary surface S of V, i.e. for $\mathbf{r}' \in S$. The connection between the function values at these two positions in space is governed by the Green's function G that belongs to the differential equation U obeys.

This theorem of Helmholtz and Kirchhoff is the central result of scalar diffraction theory on which we base our subsequent discussions. If we further impose some restrictions on U, we can manage to shrink the large surface area S to just some relevant portion of space (some small finite aperture, e.g.) and ignore the remaining part of the integral that contributes nothing to the values of $U(\mathbf{r})$, making the evaluation of the integral feasible. These restrictions will be discussed in the following.

3.3 Kirchhoff Formulation of Diffraction by a Planar Screen

We consider now the situation of Fig. 3.4, where a monochromatic plane wave is incident from the left side of an opaque screen with an aperture. We want to use (3.24) to compute the resulting wave envelope $U(\mathbf{r})$ on the right side of the aperture, where the surface $S = S_1 \cup S_2$ consists of a plane S_1 and a spherical shell S_2 , much like a bowl.

As alluded to above, the problem now is that the Helmholtz-Kirchhoff theorem requires a closed surface. It would be way more convenient, if we just specified the value of Uand its normal derivative on the screen S_1 . Hence, we have to consider the infinite halfsphere S_2 on the r.h.s of the screen and derive appropriate boundary conditions which make the contribution from this part of the surface vanish.



Figure 3.4: Kirchhoff formulation of diffraction by a plane screen. A monochromatic plane wave impinges from the left side of the screen with the aperture area Σ . We are interested in the resulting field $U(\mathbf{r})$ on the right side of the screen due to diffraction. It will be shown, that the contributions to the integral (3.24) from the sphere surface S_2 vanish.

Consider S_2 : Here's the idea: Let's make the spherical shell infinitely large, i.e., take the limit $R \to \infty$. Since the sphere surface area is of the order R^2 , we then may neglect all contributions of the integrands that decay *faster* than R^{-2} (but retain contributions of exactly the order of R^{-2}). The arguments here are motivated by physical considerations¹ as well as mathematically rigorous derivations and can be looked up also in Goodman [3]. Let's first have a look at the functions U and G that occur as integrands in (3.24). For the modulus of G we have

$$|G| = \left| \frac{\mathrm{e}^{\mathrm{i}k|\mathbf{r} - \mathbf{r}'|}}{|\mathbf{r} - \mathbf{r}'|} \right| \tag{3.25}$$

$$=\frac{1}{|\mathbf{r}-\mathbf{r}'|}\tag{3.26}$$

$$=\frac{1}{R},\tag{3.27}$$

because the modulus of the complex exponential is exactly 1 and by Fig. 3.4 the distance from \mathbf{r} is defined as the radius R of the spherical surface S_2 . G itself is an outward propagating spherical wave. From physical intuition (i.e., energy conservation) it is sensible to assume, that the wave part $U(\mathbf{r})$ also decays as fast as a spherical wave with the distance from the aperture, which is basically R when R becomes large (i.e., much larger than \mathbf{r} , the distance of the point of observation from Σ). This will additionally be justified a posteriori below. Thus, we know that the functions G and U decay as R^{-1} , which allows us to throw away those parts of the outward normal derivatives that decay faster than R^{-1} . The product will then be of the order R^{-2} and constitute the only significant contribution to the surface integral on S_2 .

¹Also known as *handwaving*.