Contents

1. Maxwell’s Equations in Free Space .......................... 3
   1.1 Coulomb’s Law and Electrostatics ..................... 3
   1.2 Special Relativity .................................. 4
       1.2.1 Space-Time Events .......................... 4
       1.2.2 The Proper Time ............................. 5
       1.2.3 Lorentz Transformation ....................... 7
       1.2.4 Dynamics of a Point Particle ................. 12
   1.3 Transformation of Electrostatic Force ............... 17
   1.4 Charge and Current Density .......................... 17
   1.5 Maxwell’s Equations ................................ 18
   1.6 Problems ........................................ 22
   1.7 Solutions ........................................ 24

2. Maxwell’s Equations in Matter ............................. 35
   2.1 The Macroscopic Maxwell’s Equations ................ 35
   2.2 The Potential 4-vector ................................ 37
   2.3 Maxwell’s Equations and Lorentz Invariance ........ 40
   2.4 Gauge Transformation ............................... 41
   2.5 The Lorenz and Coulomb Gauge Transformations in Vacuum 41
       2.5.1 Lorenz Gauge .................................. 42
       2.5.2 Coulomb Gauge ................................ 42
   2.6 Integral Representation and Boundary Conditions ...... 43
   2.7 Isotropic and Linear Medium .......................... 44
   2.8 Harmonic Time Dependency ............................ 48
   2.9 Inhomogeneous Medium Free of Sources ............... 49
   2.10 The Scalar Approximation ........................... 50
   2.11 Polarization ...................................... 50
   2.12 Problems .......................................... 54
   2.13 Solutions ......................................... 55

3. Geometrical Optics ......................................... 71
   3.1 Scalar Geometrical Optics ............................ 71
   3.2 Vectorial Geometrical Optics .......................... 72
       3.2.1 Asymptotic Expansion ......................... 73
## Contents

3.2.2 Eikonal Equation ................................ 74
3.3 Optical Rays ........................................ 75
  3.3.1 Reflection and Refraction ......................... 76
  3.3.2 The Ray Equation ................................ 77
  3.3.3 Fermat’s Principle ............................... 79
3.4 Transport Equation ...................................... 83
  3.4.1 Polarization Evolution ........................... 86
3.5 Energy Conservation .................................... 89
  3.5.1 Intensity Along an Optical Ray .................... 91
3.6 Problems ............................................... 94
3.7 Solutions ............................................. 99

4. Paraxial Approximation .................................... 117
  4.1 Paraxial Rays ....................................... 117
    4.1.1 ABCD Matrix .................................. 117
    4.1.2 Möbius Transformation ........................ 121
    4.1.3 Ray Equation .................................. 122
    4.1.4 Graded Index Medium ........................... 124
  4.2 Paraxial Waves ..................................... 125
    4.2.1 Paraxial Approximation ........................ 125
    4.2.2 Gaussian Beam in GRIN Medium ................. 125
    4.2.3 Homogeneous Case .............................. 127
  4.3 Fiber Bragg Grating ................................ 128
  4.4 Problems ........................................... 132
  4.5 Solutions .......................................... 133

5. Scalar Diffraction Theory ............................... 139
  5.1 Angular Spectrum .................................... 139
  5.2 Rayleigh-Sommerfeld, Kirchhoff, Fresnel and Fraunhofer Diffraction Integrals .................................. 144
    5.2.1 The Limit of Geometrical Optics ............... 144
  5.3 The Fresnel and Fraunhofer Diffraction Integrals .... 146
  5.4 Imaging ............................................. 148
  5.5 Problems ........................................... 151
  5.6 Solutions .......................................... 153

References .................................................. 161

Index ........................................................... 163
Preface

The laws of electromagnetic theory are expressed in this text in Gaussian units.
1. Maxwell’s Equations in Free Space

In this chapter it is shown that the Maxwell’s equations in free space can be inferred from the Coulomb’s law and the theory of special relativity.

1.1 Coulomb’s Law and Electrostatics

Consider a system containing stationary (i.e. non-moving) charged particles. Let \( \rho (\mathbf{r}') \) be the corresponding charge distribution, i.e. \( \rho (\mathbf{r}') \) is the charge per unit volume at the spacial location \( \mathbf{r}' \). A test particle having charge \( q \) is instantaneously located at the spacial point \( \mathbf{r} \). The Coulomb’s law states that the electrostatic interaction between the test particle and the charge distribution \( \rho (\mathbf{r}') \) gives rise to a force \( \mathbf{F} \) acting on the test particle that is given by

\[
\mathbf{F} = q \mathbf{E},
\]

where \( \mathbf{E} \) is the electric filed that is generated by the charge distribution \( \rho (\mathbf{r}') \). The electric field \( \mathbf{E} \) can be expressed in terms of a scalar potential \( \phi \) as

\[
\mathbf{E} = -\nabla \phi,
\]

where \( \phi \) is related to the charge distribution \( \rho (\mathbf{r}') \) by

\[
\phi (\mathbf{r}) = \int \frac{\rho (\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} \, d^3\mathbf{r}'.
\]

The expression (1.2) for the electric field \( \mathbf{E} \) implies that

\[
\nabla \times \mathbf{E} = 0.
\]

**Exercise 1.1.1.** Show that the scalar potential \( \phi \) given by Eq. (1.3) satisfies the Poisson equation, which is given by

\[
\nabla^2 \phi = -\nabla \cdot \mathbf{E} = -4\pi \rho.
\]

**Solution 1.1.1.** The Poisson equation is easily derived with the help of the general identity
Chapter 1. Maxwell’s Equations in Free Space

\[ \nabla^2 \left( \frac{1}{|\mathbf{r} - \mathbf{r}'|} \right) = -4\pi \delta(\mathbf{r} - \mathbf{r}'). \tag{1.6} \]

The above identity (1.6) can be proven by noticing that

\[ \nabla^2 \left( \frac{1}{|\mathbf{r} - \mathbf{r}'|} \right) = \nabla \cdot \left( \nabla \left( \frac{1}{|\mathbf{r} - \mathbf{r}'|} \right) \right) = \nabla \cdot \left( -\frac{\mathbf{r} - \mathbf{r}'}{|\mathbf{r} - \mathbf{r}'|^3} \right), \tag{1.7} \]

and by employing the divergence theorem [see Eq. (2.68) below] for a sphere centered at \( \mathbf{r}' \) (recall that the area of a sphere having radius \( r_s \) is \( 4\pi r_s^2 \)).

1.2 Special Relativity

In this chapter Einstein’s theory of special relativity is briefly reviewed. The transformation from the inertial frame, in which all source charges are stationary, into another inertial frame will be discussed. The transformation will be employed in order to find the relation between the above discussed electrostatic force [see Eq. (1.1)] and the corresponding force that is measured in another inertial frame [see Eq. (1.98) below].

1.2.1 Space-Time Events

Consider an event in space-time. Let \( \mathbf{r} = (x_1, x_2, x_3) \) and \( t \) be the spacial location and time coordinates, respectively, of the event as measured in a given inertial frame of reference, which is labeled as \( S \). The corresponding 4-vector \( \mathbf{X} \) is given by

\[ \mathbf{X} = (x_0, x_1, x_2, x_3)^T, \tag{1.8} \]

where \( T \) labels transpose, \( x_0 \) is given by

\[ x_0 = ct, \tag{1.9} \]

and where \( c \) is the speed of light in vacuum. In this chapter bold font is employed to denote three dimensional spacial vectors (3-vectors) and capital letters are used to denote 4-vectors. Consider an additional inertial frame of reference that is labeled as \( S' \) (see Fig. 1.1), and which is moving at a constant velocity \( c\mathbf{\beta} \) with respect to the frame \( S \) (i.e. the dimensionless 3-vector \( \mathbf{\beta} \) is the relative velocity of \( S' \) with respect to \( S \) in units of \( c \)). Let \( \mathbf{X}' = (x'_0, x'_1, x'_2, x'_3)^T \) be the 4-vector of the same event as measured in \( S' \). Consider a second event having coordinates \( \mathbf{X} + d\mathbf{X} \), where

\[ d\mathbf{X} = (dx_0, dx_1, dx_2, dx_3)^T \tag{1.10} \]

is considered as infinitesimally small. The transformed 4-vector \( d\mathbf{X}' \) is given by
1.2. Special Relativity

The inertial frames of reference $S$ and $S'$.

$$dX' = A dX,$$  \hspace{1cm} (1.11)

where the $4 \times 4$ matrix $A$ is given by

$$A = \frac{\partial (x'_0, x'_1, x'_2, x'_3)}{\partial (x_0, x_1, x_2, x_3)}. \hspace{1cm} (1.12)$$

Translational symmetry of space-time implies that $A$ is independent of $X$.

Note that, in general, a four dimensional vector is considered as a 4-vector only when it is transformed according to the Lorentz transformation, which will be discussed below.

1.2.2 The Proper Time

The proper time $d\tau$ corresponding to $dX$ is defined by

$$(d\tau)^2 \equiv c^{-2} (dX)^T \eta (dX),$$  \hspace{1cm} (1.13)

where the so-called Minkowski metric $\eta$ is given by

$$\eta = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix},$$  \hspace{1cm} (1.14)

thus

$$(d\tau)^2 = \frac{(dx_0)^2 - (dx_1)^2 - (dx_2)^2 - (dx_3)^2}{c^2} = (dt)^2 \left(1 - \frac{u \cdot u}{c^2}\right),$$  \hspace{1cm} (1.15)
where \( dt = c^{-1} dx_0 \), the velocity 3-vector \( \mathbf{u} \) is given by

\[
\mathbf{u} = \frac{d\mathbf{r}}{dt} = (u_1, u_2, u_3) = \begin{pmatrix} \frac{dx_1}{dt} \\ \frac{dx_2}{dt} \\ \frac{dx_3}{dt} \end{pmatrix},
\]

(1.16)

and \( \mathbf{u} \cdot \mathbf{u} = u_1^2 + u_2^2 + u_3^2 \).

The pair of events \( X \) and \( X + dX \) can be categorized as follows. When \( (d\tau)^2 = 0 \) (i.e. when \( c^2 (d\tau)^2 = (d\mathbf{r})^2 \)), or when \( u \equiv |\mathbf{u}| = c \) the pair of events is referred to as light-like, when \( (d\tau)^2 > 0 \) (i.e. when \( c^2 (d\tau)^2 > (d\mathbf{r})^2 \)), or when \( u < c \) as time-like and when \( (d\tau)^2 < 0 \) (i.e. when \( c^2 (d\tau)^2 < (d\mathbf{r})^2 \), or when \( u > c \) as space-like.

**Postulate** - In the theory of special relativity it is *postulated* that the speed of light \( c \) is invariant, i.e. it is assumed that the same value is measured in any inertial frame of reference. In other words, it is postulated that for the case of pair of light-like events the proper time vanishes, i.e. \( (d\tau)^2 = 0 \), in any inertial frame of reference.

**Claim.** The above postulate implies that the proper-time \( (d\tau)^2 \) is invariant (for a general type of pair of events).

**Proof.** As can be seen from Eq. (1.13), \( (d\tau)^2 \) is independent on the direction of the velocity 3-vector \( \mathbf{u} \). This property implies that the value \( (d\tau')^2 \) as being measured in an inertial frame \( S' \) having relative velocity \( \mathbf{v}' \) (with respect to the frame \( S \), in which the measured value is \( (d\tau)^2 \)) is expected to be independent on the direction of the 3-vector \( \mathbf{v}' \), i.e. \( (d\tau')^2 \) can be expressed as a function of \( (d\tau)^2 \) and \( v' = |\mathbf{v}'| \). Since the proper time is defined as infinitesimally small this function can be expressed as a linear function of \( (d\tau)^2 \)

\[
(d\tau')^2 = a_0 (v') + a_1 (v') (d\tau)^2 ,
\]

(1.17)

where both \( a_0 \) and \( a_1 \) are functions of \( v' \). The postulate that the speed of light \( c \) is invariant, i.e. the assumption that \( (d\tau')^2 = 0 \) when \( (d\tau)^2 = 0 \), implies that \( a_0 (v') = 0 \). To complete the proof one has to show that \( a_1 (v') = 1 \). This can be done by considering a third inertial frame \( S'' \) having relative velocity \( \mathbf{v}'' \) with respect to the frame \( S \). With the help of Eq. (1.17) one finds that

\[
(d\tau'')^2 = a_1 (v'') (d\tau)^2 = a_1 (v'_R) a_1 (v') (d\tau)^2 ,
\]

(1.18)

where \( v'_R = |\mathbf{v}'_R| \), and where \( \mathbf{v}'_R \) is the relative velocity of frame \( S'' \) with respect to frame \( S' \), thus the following is required to hold

\[
a_1 (v'') = a_1 (v'_R) a_1 (v') .
\]

(1.19)

The velocity \( v'_R \) is expected to depend on the angle between \( \mathbf{v}' \) and \( \mathbf{v}'' \), and thus (1.19) can hold for arbitrary 3-vectors \( \mathbf{v}' \) and \( \mathbf{v}'' \) only if \( a_1 (v) = 1 \).
1.2.3 Lorentz Transformation

The requirement that the proper-time is invariant can be expressed as [see Eqs. (1.11) and (1.13)]

$$A^T \eta A = \eta. \quad (1.20)$$

Any matrix $A$ satisfying Eq. (1.20) is called a Lorentz transformation.

**Exercise 1.2.1.** Show that $\det A = \pm 1$

**Solution 1.2.1.** In general, $\det A^T = \det A$ for any square matrix $A$ and $\det (AB) = \det (A) \det (B)$ for any pair of square matrices $A$ and $B$, thus [see Eqs. (1.14) and (1.20)] $(\det A)^2 = 1$.

**Exercise 1.2.2.** Find an expression for $A^{-1}$.

**Solution 1.2.2.** By multiplying Eq. (1.20) from the left by $\eta^{-1} = \eta$ one obtains

$$A^{-1} = \eta A^T \eta. \quad (1.21)$$

**Exercise 1.2.3.** Show that if both $A_1$ and $A_2$ are Lorentz transformations then $A_1 A_2$ is a Lorentz transformation.

**Solution 1.2.3.** With the help of Eq. (1.20) one obtains

$$(A_1 A_2)^T \eta A_1 A_2 = A_2^T A_1^T \eta A_1 A_2 = A_2^T \eta A_2 = \eta, \quad (1.22)$$

and thus $A_1 A_2$ is a Lorentz transformation.

As an example, consider the case where the inertial frame of reference $S'$ moves at a constant velocity $\beta c$ in the $x_1$ direction with respect to the frame $S$. For this case it is expected that $x'_2 = x_2$ and $x'_3 = x_3$, and consequently the transformation matrix $A$, which relates the vectors of coordinates $dX$ and $dX'$ by $dX' = AdX$ [see Eq. (1.11)], can be expressed in a block form as

$$A = \begin{pmatrix} B_1 & 0 \\ 0 & 1 \end{pmatrix}, \quad (1.23)$$

where

$$B_1 = \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix} \quad (1.24)$$

is a $2 \times 2$ matrix relating the time and $x_1$ coordinates

$$\begin{pmatrix} cdt' \\ dx'_1 \end{pmatrix} = B_1 \begin{pmatrix} cdt \\ dx_1 \end{pmatrix}. \quad (1.25)$$
Chapter 1. Maxwell’s Equations in Free Space

Claim. The $2 \times 2$ matrix $B_1$ is given by

$$B_1 = \gamma \begin{pmatrix} 1 & -\beta \\ -\beta & 1 \end{pmatrix}, \quad (1.26)$$

where

$$\gamma = \frac{1}{\sqrt{1-\beta^2}}. \quad (1.27)$$

Proof. For events occurring at $dx_1' = 0$ the following holds (since the relative velocity of $S'$ with respect to $S$ is $\beta c$)

$$\frac{1}{c} \frac{dx_1}{dt} = \beta, \quad (1.28)$$

and thus [see second row of Eq. (1.25)]

$$0 = dx_1' = b_{21} c dt + b_{22} dx_1 = \left( \frac{b_{21}}{\beta} + b_{22} \right) dx_1. \quad (1.29)$$

Similarly, for events occurring at $dx_1 = 0$ the following holds (since the relative velocity of $S$ with respect to $S'$ is $-\beta c$)

$$\frac{1}{c} \frac{dx_1'}{dt'} = -\beta, \quad (1.30)$$

and thus [see of Eq. (1.25)]

$$\frac{1}{c} \frac{dx_1'}{dt'} = \frac{b_{21}}{b_{11}} = -\beta, \quad (1.31)$$

and therefore $B_1$ can be expressed as

$$B_1 = \gamma \begin{pmatrix} 1 & b_{12} \\ -b_{12} & 1 \end{pmatrix}, \quad (1.32)$$

where $b_{11} = b_{22} \equiv \gamma$. Both unknowns $\gamma$ and $b_{12}$ can be evaluated with the help of Eq. (1.20), which yields

$$\gamma^2 \begin{pmatrix} 1-\beta^2 & \frac{b_{12}+\beta \gamma}{\gamma} \\ \frac{b_{12}+\beta \gamma}{\gamma} & -\frac{b_{12}+\gamma^2}{\gamma^2} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad (1.33)$$

and thus Eq. (1.26) holds.

Two important effects can be demonstrated using the two-dimensional Lorentz transformation (1.25):
1. **time dilation** - Consider the case where \( d\mathbf{x}_1 = 0 \). For this case \( dt = d\tau \), i.e. the time difference between the two events \( dt \) as measured in a frame \( S \) at which both events occur at the same location is the proper time \( d\tau \) [see Eq. (1.15)]. With the help of Eq. (1.25) one finds that

\[
    dt' = \gamma d\tau \geq d\tau. \tag{1.34}
\]

The time dilation factor \( \gamma \) depends on the relative velocity \( c\beta \) of frame \( S' \) with respect to \( S \)

\[
    \gamma = \frac{1}{\sqrt{1 - \beta^2}}. \tag{1.35}
\]

2. **length contraction** - Consider a rod lying along the \( x_1 \) axis. In a frame \( S \), in which the rod is at rest, the length of the rod is \( d\mathbf{x}_1 \). Let \( d\mathbf{x}_1' \) be the length of the rod as measured in a frame moving at velocity \( c\beta \) with respect to \( S \) along the \( x_1 \) axis (i.e. the relative velocity is assumed to be parallel to the axis of the rod). The measurement of \( d\mathbf{x}_1' \) in the frame \( S' \) is associated with two events having the same time, i.e. \( dt' = 0 \). With the help of Eq. (1.25) one finds that

\[
\begin{align*}
    \left( \begin{array}{c}
    cdt \\
    d\mathbf{x}_1
    \end{array} \right) &= B_1^{-1} \left( \begin{array}{c}
    0 \\
    d\mathbf{x}_1'
    \end{array} \right) \\
    &= \gamma \left( \begin{array}{c}
    1 \\
    \beta
    \end{array} \cdot \begin{array}{c}
    1 \\
    0
    \end{array} \right) \left( \begin{array}{c}
    0 \\
    d\mathbf{x}_1'
    \end{array} \right),
\end{align*}
\]

and thus

\[
    d\mathbf{x}_1' = \frac{d\mathbf{x}_1}{\gamma} \leq d\mathbf{x}_1, \tag{1.37}
\]

i.e. the length of the rod \( d\mathbf{x}_1' \) as measured in \( S' \) is smaller than the length as measured in a frame in which the rod is at rest.

The generalization of Eq. (1.23) for the case where the relative velocity of frame \( S' \) with respect to frame \( S \) can be pointing in an arbitrary direction is discussed in the following problem.

**Exercise 1.2.4.** The \( 4 \times 4 \) matrix \( B(\beta) \) is defined by

\[
    B(\beta) = \exp \left( -\kappa\hat{\beta} \cdot \Sigma \right), \tag{1.38}
\]

where \( \hat{\beta} \) is a unit vector pointing in the direction of \( \beta \) (i.e. \( \beta = \hat{\beta} \beta \) where \( \beta = |\beta| \)), the components of the matrix vector \( \Sigma = (\Sigma_1, \Sigma_2, \Sigma_3) \) are given by

\[
\begin{align*}
    \Sigma_1 &= \begin{pmatrix}
    0 & 0 & 0 & 0 \\
    0 & 1 & 0 & 0 \\
    0 & 0 & 1 & 0 \\
    0 & 0 & 0 & 1
    \end{pmatrix},
    \Sigma_2 &= \begin{pmatrix}
    0 & 0 & 0 & 0 \\
    0 & 1 & 0 & 0 \\
    0 & 0 & 1 & 0 \\
    0 & 0 & 0 & 1
    \end{pmatrix},
    \Sigma_3 &= \begin{pmatrix}
    0 & 0 & 0 & 0 \\
    0 & 0 & 0 & 0 \\
    0 & 0 & 0 & 0 \\
    0 & 0 & 0 & 1
    \end{pmatrix},
\end{align*}
\]

\[
\begin{align*}
    \Sigma_1 &= \begin{pmatrix}
    0 & 0 & 0 & 0 \\
    0 & 1 & 0 & 0 \\
    0 & 0 & 1 & 0 \\
    0 & 0 & 0 & 1
    \end{pmatrix},
    \Sigma_2 &= \begin{pmatrix}
    0 & 0 & 0 & 0 \\
    0 & 1 & 0 & 0 \\
    0 & 0 & 1 & 0 \\
    0 & 0 & 0 & 1
    \end{pmatrix},
    \Sigma_3 &= \begin{pmatrix}
    0 & 0 & 0 & 0 \\
    0 & 0 & 0 & 0 \\
    0 & 0 & 0 & 0 \\
    0 & 0 & 0 & 1
    \end{pmatrix},
\end{align*}
\]
and where the so-called rapidity $\kappa$ is given by

\[
\kappa = \tanh^{-1} \beta .
\]  

(1.40)

Show that

\[
B (\beta) = \begin{pmatrix}
\gamma & -\gamma \beta_1 & -\gamma \beta_2 & -\gamma \beta_3 \\
-\gamma \beta_1 & 1 + \frac{(\gamma-1)\beta_1^2}{\beta^2} & \frac{(\gamma-1)\beta_2\beta_1}{\beta^2} & \frac{(\gamma-1)\beta_3\beta_1}{\beta^2} \\
-\gamma \beta_2 & \frac{(\gamma-1)\beta_2\beta_1}{\beta^2} & 1 + \frac{(\gamma-1)\beta_2^2}{\beta^2} & \frac{(\gamma-1)\beta_3\beta_2}{\beta^2} \\
-\gamma \beta_3 & \frac{(\gamma-1)\beta_3\beta_1}{\beta^2} & \frac{(\gamma-1)\beta_3\beta_2}{\beta^2} & 1 + \frac{(\gamma-1)\beta_3^2}{\beta^2}
\end{pmatrix},
\]  

(1.41)

where

\[
\gamma = \frac{1}{\sqrt{1 - \beta^2}} .
\]  

(1.42)

**Solution 1.2.4.** The following holds

\[
\left( \hat{\beta} \cdot \Sigma \right)^n = \begin{cases}
B_O (\hat{\beta}) & n = 1, 3, \cdots \\
B_E (\hat{\beta}) & n = 2, 4, \cdots
\end{cases}
\]  

(1.43)

where

\[
B_O (\hat{\beta}) = \begin{pmatrix}
0 & b_1 & b_2 & b_3 \\
b_1 & 0 & 0 & 0 \\
b_2 & 0 & 0 & 0 \\
b_3 & 0 & 0 & 0
\end{pmatrix},
\]  

(1.44)

\[
B_E (\hat{\beta}) = \begin{pmatrix}
1 & 0 & 0 & 0 \\
0 & b_1 & b_2 & b_3 \\
0 & b_2b_1 & b_2^2 & b_2b_3 \\
0 & b_3b_1 & b_3b_2 & b_3^2
\end{pmatrix},
\]  

(1.45)

and where

\[
\hat{\beta} = (b_1, b_2, b_3) = \frac{1}{\beta} (\beta_1, \beta_2, \beta_3) .
\]  

(1.46)

This result together with Eq. (1.38) leads to

\[
B (\beta) = \exp \left( -\kappa \hat{\beta} \cdot \Sigma \right)
\]

\[
= -\sinh \left( \kappa \hat{\beta} \cdot \Sigma \right) + \cosh \left( \kappa \hat{\beta} \cdot \Sigma \right)
\]

\[
= 1 - \sinh (\kappa) B_O (\hat{\beta}) + (\cosh (\kappa) - 1) B_E (\hat{\beta}) ,
\]  

(1.47)

where $\mathbf{1}$ is the identity matrix and where [see Eq. (1.40)]
Combining the above results lead to Eq. (1.41). Note that the matrix $B(\beta)$ (1.41) satisfies the condition (1.20), i.e. $B(\beta)^T \eta B(\beta) = \eta$, and thus $B(\beta)$ is a Lorentz transformation.

**Exercise 1.2.5.** Consider a point particle whose 3-vector velocity as measured in an inertial frame $S$ is $v$. Calculate the 3-vector velocity of the particle $v'$ as measured in a frame $S'$ moving at a constant velocity $u$ with respect to the frame $S$.

**Solution 1.2.5.** The coordinates of frame $S$ are chosen such that the velocity $u$ is pointing in the $x_1$ direction. For this case the Lorentz transformation law of the space-time 4-vector $(c dt, dx_1, dx_2, dx_3)^T$ reads [see Eq. (1.26)]

\[
\begin{pmatrix}
    c dt' \\
    dx_1' \\
    dx_2' \\
    dx_3'
\end{pmatrix} =
\begin{pmatrix}
    \gamma & -\gamma \beta & 0 & 0 \\
    -\gamma \beta & \gamma & 0 & 0 \\
    0 & 0 & 1 & 0 \\
    0 & 0 & 0 & 1
\end{pmatrix}
\begin{pmatrix}
    c dt \\
    dx_1 \\
    dx_2 \\
    dx_3
\end{pmatrix},
\]  

or

\[
\begin{align*}
    dt' &= \gamma dt - \frac{\gamma \beta}{c} dx_1, \\
    dx_1' &= \gamma dx_1 - \gamma \beta c dt, \\
    dx_2' &= dx_2, \\
    dx_3' &= dx_3,
\end{align*}
\]  

where $\beta = u/c$ and $\gamma = 1/\sqrt{1 - \beta^2}$. Dividing Eqs. (1.52), (1.53), and (1.54) by $dt'$ yields

\[
\begin{align*}
    v_1' &= \frac{v_1 - \beta c}{1 - \frac{\beta c}{c}}, \\
    v_2' &= \frac{v_2}{\gamma \left(1 - \frac{\beta c}{c}\right)}, \\
    v_3' &= \frac{v_3}{\gamma \left(1 - \frac{\beta c}{c}\right)},
\end{align*}
\]

where

\[
\begin{align*}
    v_n &= \frac{dx_n}{dt}, \\
    v_n' &= \frac{dx_n'}{dt'},
\end{align*}
\]

thus (note that $c \beta v_1 = u \cdot v$)

\[
(v_1', v_2', v_3') = \left(\frac{v_1}{\gamma}, \frac{v_2}{\gamma}, \frac{v_3}{\gamma}\right) - \left(u - \frac{1}{\gamma} v_1, 0, 0\right) \frac{1 - \frac{v_1}{c^2}}{1 - \frac{u}{c^2}}.
\]
Chapter 1. Maxwell’s Equations in Free Space

In a vector form the above results reads (recall that \( u \) is pointing in the \( x_1 \) direction)

\[
v' = \frac{v}{\gamma} - \left[ 1 - \frac{1}{\gamma} \left( 1 - \frac{u \cdot v}{c^2} \right) \right] u,
\]

(1.60)

Consider the case of a massless particle moving at the speed of light \( v = c \hat{n} \), where \( \hat{n} \) is a unit vector. With the help of the identity \( \beta^2 = (1 + \gamma^{-1})(1 - \gamma^{-1}) \) one finds that for this case Eq. (1.60) yields

\[
\hat{n}' = \frac{\hat{n} - \gamma \left( 1 - \frac{\gamma}{\gamma + 1} \frac{u \cdot \hat{n}}{c} \right) u}{\gamma (1 - \frac{u \cdot \hat{n}}{c})},
\]

(1.61)

where \( \hat{n}' = v'/c \). The above result (1.61) is commonly called the aberration of light formula. It is straightforward to show that \( \hat{n}' \cdot \hat{n}' = 1 \), i.e. \( \hat{n}' \) is a unit vector.

1.2.4 Dynamics of a Point Particle

Assume the case where the above-discussed pair of events are two infinitesimally close points along a trajectory of a point like particle having mass \( m \). Consider an inertial frame of reference \( S \) whose velocity coincides with the instantaneous velocity of the particle (i.e. the instantaneous velocity of the particle measured in that frame vanishes). As can be seen from Eq. (1.15), \( d\tau = dt \) in that frame, which implies that the proper time \( d\tau \) is the time difference between the events as measured in \( S \). The transformation of \( dX \) into a frame \( S' \) having relative velocity \( u = c\beta \) with respect to \( S \) is given by [see Eq. (1.11)]

\[
dX' = \Lambda dX,
\]

(1.62)

where \( \Lambda = B(\beta) \) [see Eq. (1.38)].

**The Energy-Momentum 4-Vector.** The energy-momentum 4-vector \( P \) is defined by

\[
P = m \frac{dX}{d\tau}.
\]

(1.63)

In the frame \( S \), which moves with the particle, \( P \) is given by

\[
P = (mc, 0, 0, 0)^T.
\]

(1.64)

Note that in the notations employed here, the mass \( m \) is assumed to be a constant parameter, which is commonly called the rest mass. The invariance of the proper time \( d\tau \) and Eq. (1.11) together imply that \( P \) is Lorentz transformed according to
\[ P' = \Lambda P. \]  

(1.65)

In the frame \( S' \) it is given by
\[
P' = m \begin{pmatrix} \frac{dx'_0}{d\tau}, & \frac{dx'_1}{d\tau}, & \frac{dx'_2}{d\tau}, & \frac{dx'_3}{d\tau} \end{pmatrix}^T.
\]  

(1.66)

With the help of Eq. (1.34) one finds that
\[
\frac{1}{d\tau} = \frac{dt'}{d\tau} \frac{1}{dt'}. \gamma \frac{1}{dt'},
\]  

(1.67)

where
\[
\gamma = \frac{1}{\sqrt{1 - \frac{u^2}{c^2}}},
\]  

(1.68)

and where \( u = |u| = c|\beta| \), and thus \( P' \) can be expressed as
\[
P' = (p'_0, p'_1, p'_2, p'_3)^T = \left( \frac{E'}{c}, m\gamma u_1, m\gamma u_2, m\gamma u_3 \right)^T,
\]  

(1.69)

where
\[
E' = mc^2 \gamma.
\]  

(1.70)

Note that the following holds [see Eq. (1.68)]
\[
E' = mc^2 + \frac{mu^2}{2} + O \left( \beta^4 \right),
\]  

(1.71)

\[
(p'_1, p'_2, p'_3) = mu + O \left( \beta^3 \right),
\]  

(1.72)

thus in the limit where \( \beta = u/c \ll 1 \) the term \( E' \) (up to the constant \( mc^2 \)) becomes the Newtonian energy of the particle and \( (p'_1, p'_2, p'_3) \) becomes its Newtonian momentum vector.

The fact that \( P \) is transformed by a Lorentz transformation implies that the quantity \( p'^2_0 - p'^2_1 - p'^2_2 - p'^2_3 \) is invariant. With the help of Eqs. (1.64) and (1.69) one finds that
\[
m^2c^2 = \frac{E^2}{c^2} - p^2.
\]  

(1.73)

While the left hand side of (1.73) is frame independent, both energy \( E \) and momentum \( p \) of the particle are frame dependent. For massless particles Eq. (1.73) reads
\[
E = cp.
\]  

(1.74)

Consider a scattering process involving \( N_{\text{in}} \) incoming particles and \( N_{\text{out}} \) outgunning particles. The energy-momentum conservation law implies that
\[ \sum_{n=1}^{N_{\text{in}}} P_{n,\text{in}} = \sum_{n=1}^{N_{\text{out}}} P_{n,\text{out}}, \quad (1.75) \]

where \( P_{n,\text{in}} \) (\( P_{n,\text{out}} \)) are the energy-momentum 4-vectors of the incoming (outgoing) particles.

As can be seen from Eq. (1.69), the momentum of a particle that is conserved according to the theory of special relativity is given by \( m\gamma u \). Thus, contrary to the nonrelativistic version of the law of momentum conservation, in which the ratio between momentum and velocity is a frame-independent constant, in the relativistic version the ratio is the frame-dependent mass \( m\gamma \), where \( m \) is the rest mass.

**The Force 4-Vector.** The force 4-vector \( F \) is defined by

\[ F = \frac{dP}{d\tau}. \quad (1.76) \]

Similarly to the energy-momentum 4-vector \( P \), the invariance of the proper time \( d\tau \) implies that \( F \) is Lorentz transformed according to [see Eq. (1.11)]

\[ F' = \Lambda F. \quad (1.77) \]

The force 4-vector \( F \) is related to the force 3-vector \( f \), which is defined by

\[ f = \frac{dp}{dt}, \quad (1.78) \]

by [see Eqs. (1.34), (1.63) and (1.69)]

\[ F = \left( \frac{1}{c} \frac{dE}{d\tau} \right)^T = \gamma \left( \frac{1}{c} \frac{dE}{d\tau}, f \right)^T. \quad (1.79) \]

**Exercise 1.2.6.** Show that

\[ \frac{dE}{dt} = f \cdot v, \quad (1.80) \]

where \( f \) is the 3-vector force and \( v \) is the 3-vector velocity of a point particle having a mass \( m \), which is assumed to be a constant.

**Solution 1.2.6.** Consider the quantity \( P^T \eta F \), which is given by [see Eqs. (1.69) and (1.79)]

\[ P^T \eta F = m\gamma^2 \left( c \, v \right) \eta \left( \frac{1}{c} \frac{dE}{dt} \right) = m\gamma^2 \left( \frac{dE}{dt} \cdot f \cdot v \right). \quad (1.81) \]

In an inertial frame of reference \( S' \) whose velocity coincides with the instantaneous velocity of the particle the following holds \( E' = mc^2 \) [see Eq. (1.70)] and \( v' = 0 \), and thus \( (P^T \eta F)' \) vanishes (it is assumed that mass of the particle \( m \) is a constant). The fact that \( P^T \eta F \) is invariant under Lorentz transformation [see Eqs. (1.65) and (1.77)] implies that \( P^T \eta F = 0 \), and thus (1.80) holds.
Exercise 1.2.7. Consider a point particle having mass $m$ whose 3-vector force and 3-vector velocity as measured in an inertial frame $S$ are $f$ and $v$, respectively. Calculate the 3-vector force $f'$ as measured in a frame $S'$ moving at a constant velocity $u$ with respect to the frame $S$.

Solution 1.2.7. The coordinates of frame $S$ are chosen such that the velocity $u$ is pointing in the $x_1$ direction. The following holds [see Eqs. (1.26) and compare with Eqs. (1.55), (1.56) and (1.57)]

$$\frac{dE'}{dt'} = \gamma \frac{dt}{dt'} \left( \frac{dE}{dt} - \beta cv_1 \right), \quad (1.82)$$

$$f'_{1} = \gamma \frac{dt}{dt'} \left( f_1 - \frac{\beta}{c} \frac{dE}{dt} \right), \quad (1.83)$$

$$f'_{2} = \frac{dt}{dt'} f_2, \quad (1.84)$$

$$f'_{3} = \frac{dt}{dt'} f_3. \quad (1.85)$$

With the help of Eq. (1.51) one finds that

$$\frac{dt}{dt'} = \frac{1}{\gamma \left( 1 - \frac{\beta}{c} \frac{dx_1}{dt} \right)} = \frac{1}{\gamma \left( 1 - \frac{u \cdot v}{c^2} \right)}, \quad (1.86)$$

and thus [see Eq. (1.80)]

$$f'_{1} = f_1 - \frac{c}{1 - \frac{u \cdot v}{c^2}} f \cdot v, \quad (1.87)$$

$$f'_{2} = \frac{f_2}{\gamma \left( 1 - \frac{u \cdot v}{c^2} \right)}, \quad (1.88)$$

$$f'_{3} = \frac{f_3}{\gamma \left( 1 - \frac{u \cdot v}{c^2} \right)} \cdot \quad (1.89)$$

Exercise 1.2.8. In general, a given 3-vector $a$ can be decomposed as

$$a = a_{\parallel} + a_{\perp}, \quad (1.90)$$

where

$$a_{\parallel} = \frac{(u \cdot a)}{u^2} u, \quad (1.91)$$

is the component parallel to $u$ and $a_{\perp} = a - a_{\parallel}$ is the perpendicular one. Show that

$$f = f'_{\parallel} + \gamma f'_{\perp} + \gamma \frac{v \times (u \times f'_{\perp})}{c^2}. \quad (1.92)$$
Solution 1.2.8. In vectorial notation Eq. (1.87), (1.88) and (1.89) become

\[ f'_\parallel = f_\parallel - \frac{(f \cdot v) u}{1 - \frac{u \cdot v}{c^2}}, \tag{1.93} \]

\[ f'_\perp = \frac{f_\perp}{\gamma(1 - \frac{u \cdot v}{c^2})}. \tag{1.94} \]

With the help of the vector identity

\[ A \times (B \times C) = (A \cdot C) B - (A \cdot B) C, \tag{1.95} \]

one finds that

\[ (f \cdot v) u = v \times (u \times f) + (v \cdot u) f, \tag{1.96} \]

and thus Eq. (1.93) can be rewritten as (note that \( u \times f = u \times f_\perp \))

\[ f'_\parallel = f_\parallel - \frac{v \times (u \times f_\perp) + (v \cdot u) f}{1 - \frac{u \cdot v}{c^2}} \]

\[ = f_\parallel - \frac{v \times (u \times f_\perp)}{1 - \frac{u \cdot v}{c^2}} + \frac{(v \cdot u) f_\perp}{1 - \frac{u \cdot v}{c^2}} \]

\[ = f_\parallel - \gamma c^{-2} [v \times (u \times f'_\perp) + (v \cdot u) f'_\perp]. \tag{1.97} \]

The above result together with Eq. (1.94) lead to (1.92).

Exercise 1.2.9. Show that Eq. (1.92) can be rewritten as

\[ f = f_E + \frac{v \times f_B}{c}, \tag{1.98} \]

where

\[ f_E = f'_\parallel + \gamma f'_\perp, \tag{1.99} \]

and where

\[ f_B = \frac{u \times f_E}{c}. \tag{1.100} \]

Solution 1.2.9. On one hand Eqs. (1.92) and (1.99) yield

\[ f - f_E = \frac{\gamma v \times (u \times f'_\perp)}{c^2}. \tag{1.101} \]

On the other hand since \( f'_\parallel \) is parallel to \( u \) the following holds [see Eq. (1.99)]

\[ \frac{u \times f_E}{c} = \frac{\gamma u \times f'_\perp}{c}, \tag{1.102} \]

and thus Eq. (1.101) can be rewritten as

\[ f - f_E = \frac{v \times (u \times f_E)}{c}, \tag{1.103} \]

in agreement with Eq. (1.98).
1.3 Transformation of Electrostatic Force

As was discussed above, the force $F'$ acting on a point particle having charge $q$ that is generated by a stationary charge distribution $\rho'$ is given by $F' = qE'$ [see Eq. (1.1)], where the electric field $E'$ is related to the charge distribution $\rho'$ by [see Eq. (1.5)]

$$\nabla' \cdot E' = 4\pi \rho', \quad (1.104)$$

and it satisfies the following relation [see Eq. (1.4)]

$$\nabla' \times E' = 0. \quad (1.105)$$

The above-mentioned quantities that are labeled by a prime ($F'$, $\rho'$ and $E'$) are assumed to represent values measured in an inertial frame $S'$. Consider another inertial frame $S$ in which the velocity of the test particle is $v$. Let $u$ be the relative velocity of frame $S'$ with respect to $S$.

With the help of Eq. (1.98) one finds that the force on the test particle $F$ as measured in frame $S$ can be expressed as

$$F = q\left(E + \frac{v \times B}{c}\right), \quad (1.106)$$

where $E$ is given by

$$E = E'_\parallel + \gamma E'_\perp, \quad (1.107)$$

$E'_\parallel (E'_\perp)$ is the component of $E'$ parallel (perpendicular) to $u$ and $B$ is given by

$$B = \frac{u \times E}{c}. \quad (1.108)$$

Note that the charge $q$ is treated as a constant invariant under the Lorentz transformation.

1.4 Charge and Current Density

Let $\rho$ be the charge distribution and let $J$ be the current density as measured in frame $S$. In frame $S'$ it is assumed that the source charges are all stationary, and thus the current density $J'$ as measured in $S'$ vanishes. Consider an infinitesimal volume $dV'$ containing $N$ particle having charge $q$ each. In the frame $S$ the measured volume $dV$ is smaller due to length contraction $dV = \gamma^{-1}dV'$ [see Eq. (1.37)], and consequently (note that both $q$ and $N$ are required to be frame independent)

$$\rho = \gamma \rho'. \quad (1.109)$$
Chapter 1. Maxwell’s Equations in Free Space

In frame $S$ the source charges move at a constant velocity $\mathbf{u}$, and consequently the current distribution $\mathbf{J}$ is expected to be give by

$$\mathbf{J} = \gamma \rho' \mathbf{u}.$$  \hfill (1.110)

The above discussion demonstrates the fact that the current 4-vector $J$, which is defined by

$$J = (c \rho, J_1, J_2, J_3)^T,$$  \hfill (1.111)

is Lorentz transformed according to

$$J' = \Lambda J.$$  \hfill (1.112)

For an alternative definition of the current 4-vector see Eq. (1.145). Consider the quantity $\partial J$, where the 4-vector $\partial$ is defined by

$$\partial = \left( c^{-1} \frac{\partial}{\partial t}, \frac{\partial}{\partial x_1}, \frac{\partial}{\partial x_2}, \frac{\partial}{\partial x_3} \right).$$  \hfill (1.113)

As can be seen from Eq. (1.12) $\partial$ is transformed according to

$$\partial' = \partial A^{-1},$$  \hfill (1.114)

and thus $\partial J$, which is given by

$$\partial J = \frac{\partial \rho}{\partial t} + \nabla \cdot \mathbf{J},$$  \hfill (1.115)

is invariant. This implies that if charge is conserved in a given inertial frame, i.e. if the continuity equation, which is given by

$$0 = \frac{\partial \rho}{\partial t} + \nabla \cdot \mathbf{J},$$  \hfill (1.116)

holds in a given frame, then the invariance of $\partial J$ guarantees that charge is conserved in any other inertial frame.

1.5 Maxwell’s Equations

Claim. The fields $\mathbf{E}$ (1.107) and $\mathbf{B}$ (1.108), which are used to express the force $\mathbf{F}$ acting on the test particle according to Eq. (1.106), satisfy the Maxwell’s Eqs.

$$\nabla \times \mathbf{B} = \frac{4 \pi}{c} \mathbf{J} + \frac{1}{c} \frac{\partial \mathbf{E}}{\partial t},$$  \hfill (1.117)

$$\nabla \times \mathbf{E} = -\frac{1}{c} \frac{\partial \mathbf{B}}{\partial t},$$  \hfill (1.118)

$$\nabla \cdot \mathbf{E} = 4 \pi \rho,$$  \hfill (1.119)

$$\nabla \cdot \mathbf{B} = 0.$$  \hfill (1.120)
1.5. Maxwell’s Equations

Proof. As can be seen from Eq. (1.114), the following holds [see Eq. (1.26)]
\[
\frac{\partial}{\partial t} = \gamma \left( \frac{\partial}{\partial t'} - c \frac{\beta}{c} \frac{\partial}{\partial x'_1} \right),
\]
\[
\frac{\partial}{\partial x_1} = \gamma \left( \frac{\partial}{\partial x'_1} - \beta \frac{\partial}{c \partial t'} \right),
\]
\[
\frac{\partial}{\partial x_2} = \frac{\partial}{\partial x'_2},
\]
\[
\frac{\partial}{\partial x_3} = \frac{\partial}{\partial x'_3}.
\]
(1.121)

Since \( \mathbf{E} \) does not depend on \( t \) one finds that [see Eqs. (1.107), (1.122), (1.123) and (1.124)]
\[
\nabla \cdot \mathbf{E} = \frac{\partial E_1}{\partial x_1} + \frac{\partial E_2}{\partial x_2} + \frac{\partial E_3}{\partial x_3}
\]
\[
= \frac{\gamma \partial E'_1}{\partial x'_1} + \frac{\gamma \partial E'_2}{\partial x'_2} + \frac{\gamma \partial E'_3}{\partial x'_3}
\]
\[
= \gamma \nabla' \cdot \mathbf{E}'.
\]
(1.125)

The above result together with Eqs. (1.104) and (1.109) lead to Eq. (1.119)
\[
\nabla \cdot \mathbf{E} = 4 \pi \gamma \rho' = 4 \pi \rho.
\]
(1.126)

Using the vector identity
\[
\nabla \times (\mathbf{A} \times \mathbf{B}) = \mathbf{A} (\nabla \cdot \mathbf{B}) - \mathbf{B} (\nabla \cdot \mathbf{A}) + (\mathbf{B} \cdot \nabla) \mathbf{A} - (\mathbf{A} \cdot \nabla) \mathbf{B},
\]
(1.127)
one obtains (recall that \( \mathbf{u} \) is a constant vector)
\[
\nabla \times (\mathbf{u} \times \mathbf{E}) = \mathbf{u} (\nabla \cdot \mathbf{E}) - (\mathbf{u} \cdot \nabla) \mathbf{E},
\]
(1.128)
and thus [see Eqs. (1.108), (1.122) and (1.126)]
\[
\nabla \times \mathbf{B} = \frac{\mathbf{u} (\nabla \cdot \mathbf{E}) - (\mathbf{u} \cdot \nabla) \mathbf{E}}{c}
\]
\[
= \frac{4 \pi \rho}{c} \mathbf{u} - \frac{u \gamma}{c} \frac{\partial \mathbf{E}}{\partial x'_1}.
\]
(1.129)

On the other hand according to Eq. (1.121) the following holds (recall that \( c \beta = u \) and note that \( \mathbf{E} \) does not depend on \( t' \))
\[
\frac{1}{c} \frac{\partial \mathbf{E}}{\partial t} = - \frac{u \gamma}{c} \frac{\partial \mathbf{E}}{\partial x'_1}.
\]
(1.130)

The last two results together with Eqs. (1.109) and (1.110) lead to Eq. (1.117). Using the vector identity
\[ \nabla \cdot \left( \mathbf{A} \times \mathbf{B} \right) = \mathbf{B} \cdot \left( \nabla \times \mathbf{A} \right) - \mathbf{A} \cdot \left( \nabla \times \mathbf{B} \right), \quad (1.131) \]

one finds that [see Eq. (1.108)]
\[ \nabla \cdot \mathbf{B} = \frac{\nabla \cdot (\mathbf{u} \times \mathbf{E})}{c} \]
\[ = -\frac{1}{c} \left( \frac{\partial E_3}{\partial x_2} - \frac{\partial E_2}{\partial x_3} \right), \]
\[ (1.132) \]
or [see Eqs. (1.107), (1.123) and (1.124)]
\[ \nabla \cdot \mathbf{B} = -\frac{\gamma}{c} \left( \frac{\partial E'_3}{\partial x'_2} - \frac{\partial E'_2}{\partial x'_3} \right) = -\frac{\gamma}{c} \frac{\mathbf{u} \cdot (\nabla' \times \mathbf{E}')}{u}. \quad (1.133) \]

The above result together with Eq. (1.105) lead to Eq. (1.120). Finally, with the help of Eqs. (1.107), (1.108), (1.121), (1.122), (1.123) and (1.124) one obtains
\[ \nabla \times \mathbf{E} + \frac{1}{c} \frac{\partial \mathbf{B}}{\partial t} \]
\[ = \left( \frac{\partial E_3}{\partial x_2} - \frac{\partial E_2}{\partial x_3}, \frac{\partial E_1}{\partial x_3}, \frac{\partial E_3}{\partial x_1} - \frac{\partial E_1}{\partial x_2} \right) \]
\[ + \frac{1}{c} \frac{\partial \mathbf{u} \times \mathbf{E}}{\partial t} \]
\[ = \left( \frac{\gamma E'_3}{x'_2} - \frac{\gamma E'_2}{x'_3}, \frac{\gamma E'_1}{x'_3}, \frac{\gamma^2 E'_3}{x'_1} - \frac{\gamma^2 E'_2}{x'_2} \right) \]
\[ -\gamma^2 \left( \frac{u}{c} \right)^2 \left( 0, -\frac{\partial E'_3}{\partial x'_1}, \frac{\partial E'_2}{\partial x'_1} \right). \quad (1.134) \]

By subtracting the term \( \gamma (\nabla' \times \mathbf{E}') = 0 \) [see Eq. (1.105)] one obtains
\[ \nabla \times \mathbf{E} + \frac{1}{c} \frac{\partial \mathbf{B}}{\partial t} - \gamma (\nabla' \times \mathbf{E}') \]
\[ = \left( 1 - \gamma \right) \left( 0, \frac{\partial E'_1}{\partial x'_3}, -\frac{\partial E'_1}{\partial x'_2} \right) \]
\[ \left( -\gamma + \gamma^2 \left( 1 - \left( \frac{u}{c} \right)^2 \right) \right) \left( 0, -\frac{\partial E'_3}{\partial x'_1}, \frac{\partial E'_2}{\partial x'_1} \right), \quad (1.135) \]

thus [note that \( \gamma^2 \left( 1 - \left( \frac{u}{c} \right)^2 \right) = 1 \) and employ again Eq. (1.105)]
\[ \nabla \times \mathbf{E} + \frac{1}{c} \frac{\partial \mathbf{B}}{\partial t} - \gamma (\nabla' \times \mathbf{E}') = (1 - \gamma) \left( 0, \frac{\partial E'_1}{\partial x'_3} - \frac{\partial E'_3}{\partial x'_1}, \frac{\partial E'_2}{\partial x'_1} - \frac{\partial E'_1}{\partial x'_2} \right) \]
\[ = 0 , \]
\[ (1.136) \]

in agreement with Eq. (1.118).

Two comments are give below regarding the validity of the approach that has been employed above in order to infer Maxwell’s equations from electrostatics and special relativity.

1. The derivation above is based on the assumption that the laws of electrostatics hold (Coulomb’s law). By performing a Lorentz transformation from a given inertial frame, in which the source charges are at rest, to another inertial frame, one can infer what are the forces generated by charges moving at a constant velocity. However, this approach cannot be used to treat the question of what forces are generated by accelerating charges, since the theory of special relativity deals only with transformations between inertial frames. It is known that Maxwell’s equations are valid for the general case, in which source charges are allowed to accelerate. However, this fact cannot be inferred based on electrostatics and special relativity only.

2. Apparently, an approach similar to the one discussed in this chapter can be employed for the case of gravitational forces, starting from the assumption that the laws of ‘gravitostatics’ hold (Newton’s laws). However, the two cases are not equivalent. While the electric charge of a particle is assumed to be a constant, its mass, as is measured by a given observer, depends on the velocity of the observer (see the discussion above on relativistic momentum conservation). An alternative way to understand the difference between these two cases is related to the fact that the inertial mass (which appears in Newton’s second law \( F = ma \)) equals the gravitational mass (which appears in Newton’s law of gravitation \( F = Gm_1m_2/r^2 \)) for the attraction force \( F \) between two point particles having masses \( m_1 \) and \( m_2 \), where \( G \) is Newton’s constant and \( r \) is the distance between the particles. This fact implies that different particles having different mass fall at the same acceleration under gravitation, as was first found by Galileo Galilei. On the other hand, different particles having different charge in an electrostatic field need not fall at the same acceleration.
1.6 Problems

1. Consider three inertial frames $S$, $S'$ and $S''$. The relative velocity of $S'$ with respect to $S$ is $c\beta \hat{\beta}$ and the relative velocity of $S''$ with respect to $S'$ is $c\beta \hat{\beta}$. Find a Lorentz transformation mapping from $S$ to $S''$.

2. **The Compton effect** - Consider a photon having energy $E_{p,\text{in}}$ hitting an electron at rest. Calculate the energy of the reflected photon $E_{p,\text{out}}$ for the case where the scattered photon is back-reflected, i.e. its direction is reversed in the process.

3. Show that free electrons can neither emit nor absorb photons.

4. **The Doppler effect** - Consider a plane wave (not necessarily an electromagnetic wave) having the form

$$\psi = \psi_0 \cos \phi,$$  \hspace{1cm} (1.137)

where the amplitude $\psi_0$ is a constant, the phase $\phi$ is given by

$$\phi = \mathbf{k} \cdot \mathbf{r} - \omega t,$$  \hspace{1cm} (1.138)

where both wave 3-vector $\mathbf{k} = (k_1, k_2, k_3)$ and angular frequency $\omega$ are constants. While the values $\mathbf{k}$ and $\omega$ are measured in an inertial frame $S$, the values $\mathbf{k}'$ and $\omega'$ are measured in another inertial frame $S'$ moving at velocity $c\beta$ with respect to $S$. Calculate $\mathbf{k}'$ and $\omega'$.

5. Consider a point particle moving along a straight line (which is taken to be the $x_1$ axis) with a constant proper acceleration $a$ (the proper acceleration is defined as the acceleration in an inertial frame, commoving with the particle, in which it is instantaneously at rest). Let $S$ be a fixed inertial frame in which both the particle’s position 3-vector $\mathbf{r}(t) = (x_1(t), 0, 0)$ and velocity 3-vector $\mathbf{v}(t) = (v_1, 0, 0)$, where $v_1 = d x_1 / dt$, are assumed to vanish at time $t = 0$. Calculate $x_1(t)$.

6. Consider an observer moving along a given trajectory, which in a given inertial frame $S$ is taken to be given by $(ct_T(\tau), x_T(\tau), 0, 0)$, where $\tau$ is the proper time, i.e. $\tau$ is the time as being measured by a clock that is carried along with the moving observer. Consider a point object, whose spatial location is $(x, 0, 0)$ in the inertial frame of reference $S$. The observer sent a light signal towards the object at time $t'_1$. The light signal is reflected by the object, and it returns to the observer at a later time $t'_2$. Both $t'_1$ and $t'_2$ are measured by the clock commoving with the observer. Let $E$ denotes the event of light signal hitting the object (and reflected off the object), and let $t$ and $x$ be the coordinates of $E$ in the inertial frame of reference $S$. The moving observer assigns his own coordinates $t'$ and $x'$ to the event $E$ by employing the following relations

$$t' = \frac{t'_1 + t'_2}{2},$$  \hspace{1cm} (1.139)

$$x' = c \frac{t'_2 - t'_1}{2}.$$  \hspace{1cm} (1.140)
1.6. Problems

a) Derive relations between the coordinates $t$ and $x$ and the coordinates $t'$ and $x'$ given by Eqs. (1.139) and (1.140).
b) Simplify the derived relations for the case of a stationary observer located at the origin of the inertial frame of reference $S$.
c) The same for the case of an observer moving at a constant velocity $\beta c$ along the $x_1$ axis. Assume that at $\tau = 0$ the spatial location of the observer is $(0, 0, 0)$ in the inertial frame of reference $S$.
d) The same for the case of an observer moving at a constant proper acceleration $a$ along the $x_1$ axis. Assume that at $\tau = 0$ the velocity of the observer vanishes in $S$, and the spatial location of the observer is $(0, 0, 0)$ at $\tau = 0$ in $S$.
e) **slow observer** - The normalized observer’s velocity is denoted by

$$\beta = \frac{1}{c} \frac{dx_T}{dt_T}. \quad (1.141)$$

Show that

$$\frac{t - t'}{t} = \hat{\beta} \frac{(x - x_T(0))}{ct} + O(\beta^2), \quad (1.142)$$

where $\hat{\beta}$, which is given by

$$\hat{\beta} = \int_{t' = t}^{t' + \frac{x - x_T(0)}{c}} \frac{d\tau'}{2 \frac{x - x_T(0)}{c}}, \quad (1.143)$$

represents the averaged and normalized velocity over the time interval $[t - (x - x_T(0))/c, t + (x - x_T(0))/c]$.

7. **The Unruh-Davies Effect** - Consider an electromagnetic plane wave having amplitude $A$ and wavelength $\lambda$. The plane wave propagates along the $x_1$ axis. Let $f(t')$ be the time-dependent amplitude of the plane wave as being measured by an observer moving at a constant proper acceleration $a$ along the $x_1$ axis, where $t'$ is the time coordinate of the observer. Calculate $|f(\omega')|^2$, where $f(\omega')$ is the Fourier transform of $f(t')$, i.e.

$$f(\omega') = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} dt' \ f(t') e^{i\omega' t'}. \quad (1.144)$$

8. **current 4-vector** - Consider a medium containing point particles labeled by the index $n$. Let $q_n$ be the charge of the $n$th particle, and let $X_n(\tau_n) = (ct_n(\tau_n), r_n(\tau_n))^T$ be the trajectory in space-time of the $n$th point particle, where $\tau_n$ is the proper time, i.e. $\tau_n$ is the time as being measured by a clock that is carried along with the $n$th particle. The current 4-vector $J(X) = (cp(X), J(X))^T$ at space-time point $X = (ct, r)^T$ is defined by
Chapter 1. Maxwell’s Equations in Free Space

\[ J(X) = \sum_n J_n(X) \]  
\[ \text{where } J_n(X) = (c\rho_n, J_n)^T, \]  
which represents the contribution of the \( n \)’th particle to the total current 4-vector, is taken to be given by

\[ J_n(X) = q_n c \int d\tau_n \frac{dX_n}{d\tau_n} \delta(X - X_n(\tau_n)). \]  

Note that the invariance of the proper time \( \tau_n \) and the fact that \( dX_n \) is a 4-vector imply that \( J(X) \) is a 4-vector. Find expressions for \( \rho(X) \) and \( J(X) \).

1.7 Solutions

1. The desired transformation is given by [see Eqs. (1.38) and (1.40)]

\[ \Lambda = \exp \left( - (\kappa_1 + \kappa_2) \hat{\beta} \cdot \Sigma \right), \]  
\[ \text{where } \kappa_{1,2} = \tanh^{-1} \beta_{1,2}. \]  
Using the identity

\[ \tanh (\kappa_1 + \kappa_2) = \frac{\tanh (\kappa_1) + \tanh (\kappa_2)}{1 + \tanh (\kappa_1) \tanh (\kappa_2)}, \]  
on one finds that

\[ \beta = \frac{\beta_1 + \beta_2}{1 + \beta_1 \beta_2}, \]  
\[ \text{where } \beta = \tanh (\kappa_1 + \kappa_2). \]  

2. Energy-momentum conservation implies [see Eqs. (1.73), (1.74) and (1.75)]

\[ E_{p,\text{in}} + m_e c^2 = E_{p,\text{out}} + \sqrt{m_e^2 c^4 + p^2 c^2}, \]  
\[ \frac{E_{p,\text{in}}}{c} = \frac{E_{p,\text{out}}}{c} + p, \]  
where \( m_e \) is the electron mass and \( p \) is the momentum of the scattered electron. By solving for the unknowns \( E_{p,\text{out}} \) and \( p \) one obtains

\[ \frac{E_{p,\text{out}}}{m_e c^2} = \frac{E_{p,\text{in}}}{m_e c^2} \left( 1 + \frac{2 \frac{E_{p,\text{in}}}{m_e c^2}}{1 + \frac{2 \frac{E_{p,\text{in}}}{m_e c^2}}{\sqrt{1 + \frac{2 \frac{E_{p,\text{in}}}{m_e c^2}}{m_e c^2}}} \right). \]
3. Consider a reference frame, in which the electron is initially at rest. Energy-momentum conservation for the case of photon absorption implies that [see Eqs. (1.73), (1.74) and (1.75)]

\[
p_p = p_e, \quad (1.155)
\]

\[
p_p c = \sqrt{m_e^2 c^4 + p_e^2 c^2}, \quad (1.156)
\]

and for the case of photon emission that

\[
0 = p_p + p_e, \quad (1.157)
\]

\[
0 = p_p c + \sqrt{m_e^2 c^4 + p_e^2 c^2}, \quad (1.158)
\]

where \( p_p \) and \( p_e \) denote the momentum 3-vector of the photon and electron, respectively, and \( m_e \) is the electron mass. Clearly, for both cases the only possible solution is \( p_p = p_e = 0 \).

4. Consider the 4-vector \( K \), which is defined by

\[
K = -\partial \eta \phi = \left( \frac{\omega}{c}, k_1, k_2, k_3 \right), \quad (1.159)
\]

where \( \partial = \left( c^{-1} \partial/\partial t, \partial/\partial x_1, \partial/\partial x_2, \partial/\partial x_3 \right) \) [see Eq. (1.113)] and where \( \phi = k \cdot \mathbf{r} - \omega t \) [see Eq. (1.138)]. Since \( \partial \) is transformed according to \( \partial' = \partial \Lambda^{-1} \) [see Eq. (1.114)] and because \( \phi \) is expected to be Lorentz invariant (explain why) one concludes that the 4-vector \( K \) is transformed according to [see Eq. (1.21)]

\[
K' = -\partial' \eta \phi = -\partial \Lambda^{-1} \eta \phi = -\partial \eta \Lambda^T \phi. \quad (1.160)
\]

The coordinates of frame \( S \) are chosen such that the velocity \( c \beta \) is pointing in the \( x_1 \) direction. For this case Eq. (1.160) becomes [see Eqs. (1.14) and (1.26)]

\[
K' = \left( \gamma \left( -\frac{1}{c} \frac{\partial \phi}{\partial t} - \beta \frac{\partial \phi}{\partial x_1} \right), \gamma \left( \frac{\beta}{c} \frac{\partial \phi}{\partial t} + \frac{\partial \phi}{\partial x_1} \right), \frac{\partial \phi}{\partial x_2}, \frac{\partial \phi}{\partial x_3} \right),
\]

\[
= \left( \gamma \left( \frac{\omega}{c} - \beta k_1 \right), \gamma \left( \frac{\beta \omega}{c} + k_1 \right), k_2, k_3 \right), \quad (1.161)
\]

where \( \beta = |\beta| \) and \( \gamma = 1/\sqrt{1 - \beta^2} \), thus

\[
\omega' = \frac{\gamma \omega}{1 - \frac{c \beta k_1}{\omega}}, \quad (1.162)
\]

and

\[
k' = \left( \gamma \left( k_1 - \frac{\beta \omega}{c} \right), k_2, k_3 \right). \quad (1.163)
\]

Let \( \theta (\theta') \) be the angle between \( k \) (\( k' \)) and \( \beta \).
\[ \theta = \cos^{-1} \frac{k_1}{k}, \]  
(1.164)  
\[ \theta' = \cos^{-1} \frac{k'_1}{k'}, \]  
(1.165)  

where  
\[ k = \sqrt{k_1^2 + k_2^2 + k_3^2}, \]  
(1.166)  
\[ k' = \sqrt{k_1'^2 + k_2'^2 + k_3'^2}. \]  
(1.167)  

Using this notation Eq. (1.162) can be rewritten as  
\[ \frac{\omega'}{\omega} = \frac{1 - \frac{ck}{\omega} \beta \cos \theta}{\sqrt{1 - \beta^2}}. \]  
(1.168)  

The angle \( \theta' \) can be found from Eq. (1.163)  
\[ \tan \theta' = \frac{\sin \theta \sqrt{1 - \beta^2}}{\cos \theta - \frac{\omega}{ck}}. \]  
(1.169)  

Note that for case of an electromagnetic wave \( ck = \omega \).  

5. In the instantaneous rest frame of the particle the velocity 4-vector of the particle \( U \) is given by [see Eq. (1.63)]  
\[ U = \frac{dX}{d\tau} = (c, 0, 0, 0). \]  
(1.170)  

and the acceleration 4-vector \( A \) is given by  
\[ A = \frac{d^2X}{d\tau^2} = (0, a, 0, 0). \]  
(1.171)  

Using the Lorentz transformation for \( U \) and \( A \) one obtains  
\[ \frac{d}{d\tau} \begin{pmatrix} ct \\ x_1 \end{pmatrix} = B_1 \begin{pmatrix} c \\ 0 \end{pmatrix}, \]  
(1.172)  
and  
\[ \frac{d^2}{d\tau^2} \begin{pmatrix} ct \\ x_1 \end{pmatrix} = B_1 \begin{pmatrix} 0 \\ a \end{pmatrix}, \]  
(1.173)  

where [see Eq. (1.26)]  
\[ B_1 = \gamma \begin{pmatrix} 1 & -\beta \\ -\beta & 1 \end{pmatrix}, \]  
(1.174)  
\[ \beta = -v_1/c \] and \( \gamma = 1/\sqrt{1 - v_1^2/c^2} \). With the help of Eq. (1.174) Eq. (1.172) becomes
\[
\left( \frac{dt}{d\tau} \right) = \left( \frac{\gamma}{v_1} \right) .
\]  
(1.175)

Substituting Eq. (1.172) into Eq. (1.173) yields
\[
\left( \frac{d\gamma}{dt \gamma v_1} \right) = \left( \frac{v_1 a}{a} \right) .
\]  
(1.176)

The second equation of (1.176), which reads
\[
d \left( \frac{v_1}{\sqrt{1 - v_1^2/c^2}} \right) = a ,
\]  
(1.177)
can be solved using the transformation
\[
\frac{v_1}{c} = \tanh s ,
\]  
(1.178)
which yields (recall that \(1 - \tanh^2 s = 1/\cosh^2 s\))
\[
d \left( \sinh s \right) = \frac{a}{c} .
\]  
(1.179)

Integration and employing the initial conditions at time \(t = 0\) lead to [see Eq. (1.178)]
\[
\frac{v_1}{\sqrt{1 - v_1^2/c^2}} = at ,
\]  
(1.180)
and thus
\[
v_1 (t) = \frac{at}{\sqrt{1 + \left( \frac{at}{c} \right)^2}} .
\]  
(1.181)

By integrating Eq. (1.181) one obtains
\[
x_1 (t) = \frac{c^2}{a} \left( \sqrt{1 + \frac{a^2 t^2}{c^2}} - 1 \right) .
\]  
(1.182)

Alternatively, the position \(x_1\) and velocity \(v_1\) can be expressed as a function of the proper time \(\tau\) as follows. With the help of Eq. (1.181) the first equation of (1.175) becomes
\[
\frac{dt}{d\tau} = \sqrt{1 + \left( \frac{at}{c} \right)^2} ,
\]  
(1.183)
and thus
Chapter 1. Maxwell’s Equations in Free Space

\[ t = \frac{c}{a} \sinh \frac{a\tau}{c}. \] (1.184)

The last result (1.184) together with Eqs. (1.181) and (1.182) yield

\[ v_1(\tau) = c \tanh \frac{a\tau}{c}, \] (1.185)

and

\[ x_1(\tau) = \frac{c^2}{a} \left( \cosh \frac{a\tau}{c} - 1 \right). \] (1.186)

6. The light signal sent by the observer connects the space-time points \((ct, x_T(\tau) (t_1'), x_T(\tau) (t_1'), 0, 0)\) and \((ct, x, 0, 0)\), and the back reflected light signal connects the space-time points \((ct, x, 0, 0)\) and \((ct, x_T(\tau) (t_2'), x_T(\tau) (t_2'), 0, 0)\), and thus

\[ \frac{x - x_T(\tau) (t_1')}{t - t_T(\tau) (t_1')} = c, \] (1.187)

\[ \frac{x_T(\tau) (t_2') - x}{t_T(\tau) (t_2') - t} = -c, \] (1.188)

a) With the help of Eqs. (1.139) and (1.140) the relations (1.187) and (1.188) can be rewritten as

\[ \frac{x - x_T(\tau) (t' - \frac{x'}{c})}{t - t_T(\tau) (t' - \frac{x'}{c})} = c, \] (1.189)

\[ \frac{x_T(\tau) (t' + \frac{x'}{c}) - x}{t_T(\tau) (t' + \frac{x'}{c}) - t} = -c, \] (1.190)


or

\[ x - ct = x_T(\tau) \left( t' - \frac{x'}{c} \right) - ct_T(\tau) \left( t' - \frac{x'}{c} \right), \] (1.191)

\[ x + ct = x_T(\tau) \left( t' + \frac{x'}{c} \right) + ct_T(\tau) \left( t' + \frac{x'}{c} \right). \] (1.192)

b) For this case \(t_T(\tau) = \tau\) and \(x_T(\tau) = 0\), and thus Eqs. (1.191) and (1.192) become

\[ x - ct = -c \left( t' - \frac{x'}{c} \right), \] (1.193)

\[ x + ct = c \left( t' + \frac{x'}{c} \right). \] (1.194)

The solution is \(t' = t\) and \(x' = x\).

c) For this case \(t_T(\tau) = \gamma \tau\) and \(x_T(\tau) = \beta ct = \beta c\gamma \tau\) [see Eq. (1.34)], where

\[ \gamma = \frac{1}{\sqrt{1 - \beta^2}}, \] (1.195)
and thus Eqs. (1.191) and (1.192) become
\[ x - ct = \beta c \gamma \left( t' - \frac{x'}{c} \right) - c \gamma \left( t' - \frac{x'}{c} \right), \quad (1.196) \]
\[ x + ct = \beta c \gamma \left( t' + \frac{x'}{c} \right) + c \gamma \left( t' + \frac{x'}{c} \right). \quad (1.197) \]
The solution can be written as [compare with Eq. (1.25)]
\[ \begin{pmatrix} ct' \\ x' \end{pmatrix} = \gamma \begin{pmatrix} 1 & -\beta \\ -\beta & 1 \end{pmatrix} \begin{pmatrix} ct \\ x \end{pmatrix}. \quad (1.198) \]
d) For this case [see Eqs. (1.184) and (1.186)]
\[ t_T(\tau) = \frac{c^2}{a} \sinh \frac{a \tau}{c}, \quad (1.199) \]
\[ x_T(\tau) = \frac{c^2}{a} \left( \cosh \frac{a \tau}{c} - 1 \right), \quad (1.200) \]
and thus Eqs. (1.191) and (1.192) become
\[ x - ct = \frac{c^2}{a} \left( e^{\frac{a(x'-t')}{c}} - 1 \right), \quad (1.201) \]
\[ x + ct = \frac{c^2}{a} \left( e^{\frac{a(x'+t')}{c}} - 1 \right). \quad (1.202) \]
The solution is given by
\[ x = \frac{c^2}{a} \left[ \exp \left( \frac{ax'}{c^2} \right) \cosh \left( \frac{at'}{c} \right) - 1 \right], \quad (1.203) \]
\[ ct = \frac{c^2}{a} \exp \left( \frac{ax'}{c^2} \right) \sinh \left( \frac{at'}{c} \right). \quad (1.204) \]
To lowest nonvanishing order in \( a \)
\[ x = \left( 1 + \frac{x'}{2 \pi} \right) x' + \frac{at'^2}{2} + O \left( a^2 \right), \quad (1.205) \]
\[ ct = ct' \left( 1 + \frac{x'}{2 \pi} \right) + O \left( a^2 \right). \quad (1.206) \]
The inverse transformation is given by
\[ x' = \frac{c^2}{2a} \log \left[ \left( 1 + \frac{x}{c^2} \right)^2 - \left( \frac{at}{c} \right)^2 \right] \]
\[ = \left( 1 - \frac{x}{2 \pi} \right) x - \frac{at^2}{2} + O \left( a^2 \right), \quad (1.207) \]
\[ ct' = \frac{c^2}{a} \tanh^{-1} \left( \frac{ct}{x + \frac{c^2}{a}} \right) = ct \left( 1 - \frac{x}{\frac{c^2}{a}} \right) + O \left( a^2 \right). \]  

(1.208)

e) The transformation of a velocity 2-vector between \( S \) and the instantaneous rest frame of the observer is given by

\[ \frac{d}{d\tau} \left( \begin{array}{c} ct_T \\ x_T \end{array} \right) = A(-\beta) \left( \begin{array}{c} c \\ 0 \end{array} \right), \]  

(1.209)

where \( A(-\beta) \), which is given by [see Eq. (1.25)]

\[ A(-\beta) = \gamma \left( \begin{array}{cc} 1 & \beta \\ \beta & 1 \end{array} \right), \]  

(1.210)

is the 1 + 1 dimensional Lorentz transformation, \( \beta = v/c, \gamma = 1/\sqrt{1 - v^2/c^2} \) and

\[ v = \frac{dx_T}{dt_T}. \]  

(1.211)

Integrating Eq. (1.209), which can be rewritten as

\[ \frac{dt_T}{d\tau} = \gamma, \]  

(1.212)

\[ \frac{dx_T}{d\tau} = c\gamma\beta, \]  

(1.213)

with the assumed initial condition \( t_T (\tau = 0) = 0 \), yields

\[ t_T (\tau) = \int_0^\tau d\tau' \gamma (\tau') , \]  

(1.214)

\[ x_T (\tau) = x_T (0) + c \int_0^\tau d\tau' \gamma (\tau') \beta (\tau') . \]  

(1.215)

With the help of Eqs. (1.214) and (1.215) one finds that

\[ x_T (\tau) \pm ct_T (\tau) = x_T (0) + c \int_0^\tau d\tau' \gamma (\beta \pm 1) , \]  

(1.216)

and thus Eqs. (1.191) and (1.192) can be rewritten as

\[ \frac{x - x_T (0)}{c} - t = \int_0^{t - \frac{x'}{c}} d\tau' \gamma (\beta - 1) , \]  

(1.217)

\[ \frac{x - x_T (0)}{c} + t = \int_0^{t + \frac{x'}{c}} d\tau' \gamma (\beta + 1) , \]  

(1.218)
or alternatively, with the help of the relations

\[ -\gamma + \beta \gamma = -\sqrt{\frac{1-\beta}{1+\beta}}, \quad (1.219) \]

\[ -\gamma - \beta \gamma = -\sqrt{\frac{1+\beta}{1-\beta}}, \quad (1.220) \]

as

\[
t' - x' = t - \frac{x - x_T(0)}{c} + \int_0^{t'} \frac{\sqrt{1-\beta}}{1+\beta} \, dt' \left( 1 - \sqrt{\frac{1-\beta}{1+\beta}} \right), \quad (1.221) \\
t' + x' = t + \frac{x - x_T(0)}{c} + \int_0^{t'} \frac{\sqrt{1+\beta}}{1-\beta} \, dt' \left( 1 - \sqrt{\frac{1+\beta}{1-\beta}} \right). \quad (1.222) 
\]

In the limit of a stationary observer having a vanishing velocity both integrals on the right hand sides of Eqs. (1.221) and (1.222) vanish. The solutions of Eqs. (1.221) and (1.222) in this limit provides approximated values for the upper limits of these integrals, which can be used to turn Eqs. (1.221) and (1.222) into

\[
t' - x' = t - \frac{x - x_T(0)}{c} + \int_0^{t'} \frac{\sqrt{1-\beta}}{1+\beta} \, dt' \left( 1 - \sqrt{\frac{1-\beta}{1+\beta}} \right) + O(\beta^2), \quad (1.223) \\
t' + x' = t + \frac{x - x_T(0)}{c} + \int_0^{t'} \frac{\sqrt{1+\beta}}{1-\beta} \, dt' \left( 1 - \sqrt{\frac{1+\beta}{1-\beta}} \right) + O(\beta^2). \quad (1.224) 
\]

By adding the above equations one obtains

\[
\frac{t-t'}{t} = -1 + \frac{1}{2t} \int_0^{t'} \frac{\sqrt{1-\beta}}{1+\beta} \, dt' \left( 1 - \sqrt{\frac{1-\beta}{1+\beta}} \right) + \frac{1}{2t} \int_0^{t'} \frac{\sqrt{1+\beta}}{1-\beta} \, dt' \left( 1 - \sqrt{\frac{1+\beta}{1-\beta}} \right). \quad (1.225) 
\]
Chapter 1. Maxwell’s Equations in Free Space

The expansions
\[
\sqrt{\frac{1-\beta}{1+\beta}} = 1 - \beta + O(\beta^2),
\]
\[
\sqrt{\frac{1+\beta}{1-\beta}} = 1 + \beta + O(\beta^2),
\]
lead to Eq. (1.142). Note that for the case of a constant \(\beta\) Eq. (1.142) yields
\[
\frac{t-t'}{t} = \frac{\beta(x-x_T(0))}{ct},
\]
whereas the exact result (1.25) is
\[
\frac{t-t'}{t} = \frac{\gamma \beta(x-x_T(0))}{ct} + 1 - \gamma,
\]
in agreement with the approximated result only to first order in \(\beta\).

7. The plane wave can be expressed using either the inertial frame coordinates \(t\) and \(x_1\)
\[
f(t, x_1) = Ae^{i\omega_0\left(\frac{x_1}{c} - t\right)},
\]
where
\[
\omega_0 = \frac{2\pi c}{\lambda},
\]
or the observer’s coordinates \(t'\) and \(x'_1\) [see Eq. (1.201)]
\[
f(t', x'_1) = A \exp\left(\frac{i\omega_0 c}{a} \left(\frac{x_1}{c} - t'\right) - a\left(\frac{x_1}{c} - t\right)\right)
\]  
\[
= Ae^{-\frac{i\omega_0}{\omega_a}} e^{iae^{-\omega_a t'}},
\]
where
\[
\omega_a = \frac{a e^{-\omega_0 t'}}{c},
\]
and where
\[
\alpha = \frac{\omega_0}{\omega_a} e^{\omega_a x'_1}.
\]
By employing the integral variable transformation \(z = \alpha e^{-\omega_0 t'}\) one obtains [see Eq. (1.144)]
1.7. Solutions

\[ f (\omega') = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} dt' \ f (t', x') e^{i\omega't'} = Ae^{-\frac{i\omega_0}{\alpha} \frac{i\omega'}{\omega_n}} \frac{1}{\sqrt{2\pi\omega_n}} \int_0^\infty dz \ e^{iz^2 - \frac{i\omega'}{\omega_n}} , \]  

(1.234)

and thus

\[ |f (\omega')|^2 = \frac{|A|^2}{\omega_n\omega'} n_{\text{BE}} \left( \frac{2\pi\omega'}{\omega_n} \right) , \]  

(1.235)

where

\[ n_{\text{BE}} (\epsilon) = \frac{1}{e^\epsilon - 1} \]  

(1.236)

is the Bose–Einstein distribution function. In terms of the so-called Unruh-Davies temperature \( T_{\text{UD}} \), which is given by

\[ T_{\text{UD}} = \frac{\hbar a}{2\pi k_B c} , \]  

(1.237)

where \( k_B \) is the Boltzmann’s constant, the result can be rewritten as

\[ |f (\omega')|^2 = \frac{|A|^2}{\omega_n\omega'} n_{\text{BE}} \left( \frac{\hbar a}{k_B T_{\text{UD}}} \right) . \]  

(1.238)

8. With the help of Eqs. (1.145) and (1.146) one obtains

\[ \rho (t, r) = \sum_n q_n \int d\tau_n \frac{dt_n}{d\tau_n} \delta (t - t_n (\tau_n)) \delta (r - r_n (\tau_n)) , \]  

(1.239)

and

\[ J (t, r) = \sum_n q_n \int d\tau_n \frac{dr_n}{d\tau_n} \delta (t - t_n (\tau_n)) \delta (r - r_n (\tau_n)) , \]  

(1.240)

thus with the help of the relation

\[ \delta (t - t_n (\tau_n)) = \left( \left| \frac{dt_n}{d\tau_n} \right| \right)^{-1} \delta (\tau_n - \tau_n (t_n)) , \]  

(1.241)

one finds that

\[ \rho (t, r) = \sum_n q_n \delta (r - r_n (t)) , \]  

(1.242)

and

\[ J (t, r) = \sum_n q_n v_n (t) \delta (r - r_n (t)) , \]  

(1.243)

where \( r_n (t) \) and \( v_n (t) \) are the location and velocity, respectively, of the \( n \)'th particle at time \( t \).
2. Maxwell’s Equations in Matter

In this chapter the macroscopic Maxwell’s equations, which are used to described electromagnetic fields in matter, are derived.

2.1 The Macroscopic Maxwell’s Equations

When dealing with electromagnetic fields inside matter it is convenient to replace the fields $\mathbf{E}$ and $\mathbf{B}$ and sources $\phi$ and $\mathbf{J}$ appearing in the Maxwell’s equations (1.117), (1.118), (1.119) and (1.120) by their spacial average according to the following procedure

$$\psi(\mathbf{r}) \rightarrow \langle \psi(\mathbf{r}) \rangle = \frac{1}{\Delta V} \int_{\Delta V} d\mathbf{r}' \psi(\mathbf{r} + \mathbf{r}') , \quad (2.1)$$

where $\Delta V$ is the averaging volume, which is chosen such that $d_A \ll \Delta V^{1/3} \ll \lambda$, where $d_A$ is the characteristic distance between atoms in the matter and where $\lambda$ is the characteristic wavelength of electromagnetic fields. The averaged fields and sources satisfy the same set of Maxwell’s equations (1.117), (1.118), (1.119) and (1.120)

$$\nabla \times \langle \mathbf{B} \rangle = \frac{4\pi}{c} \langle \mathbf{J}_{\text{total}} \rangle + \frac{1}{c} \frac{\partial \langle \mathbf{E} \rangle}{\partial t} , \quad (2.2)$$

$$\nabla \times \langle \mathbf{E} \rangle = -\frac{1}{c} \frac{\partial \langle \mathbf{B} \rangle}{\partial t} , \quad (2.3)$$

$$\nabla \cdot \langle \mathbf{E} \rangle = 4\pi \langle \rho_{\text{total}} \rangle , \quad (2.4)$$

$$\nabla \cdot \langle \mathbf{B} \rangle = 0 . \quad (2.5)$$

Note that the label 'total' has been added as a subscript to $\rho$ and $\mathbf{J}$. This is done because in what follows the charge density $\rho = \rho_{\text{total}}$ and current density $\mathbf{J} = \mathbf{J}_{\text{total}}$ are both decomposed into different parts. To avoid cumbersome notation, the averaging symbol $\langle \rangle$ is henceforth omitted.

Electromagnetic fields in matter may result in dielectric polarization $\mathbf{P}$ and magnetization $\mathbf{M}$. It is convenient to decompose the charge and current densities into parts associated with dielectric polarization and magnetization and parts associated with other contributions. The total charge density $\rho_{\text{total}}$ is decomposed as
\[ \rho_{\text{total}} = \rho_{\text{ext}} + \rho_{\text{pol}}, \quad (2.6) \]

where \( \rho_{\text{pol}} \), which represents the contribution due to dielectric polarization, is given by

\[ \rho_{\text{pol}} = -\nabla \cdot \mathbf{P}, \quad (2.7) \]

and \( \rho_{\text{ext}} \) represents all other contributions. To total current density \( \mathbf{J}_{\text{total}} \) is decomposed as

\[ \mathbf{J}_{\text{total}} = \mathbf{J}_{\text{cond}} + \mathbf{J}_{\text{bound}} + \mathbf{J}_{\text{ext}}, \quad (2.8) \]

where \( \mathbf{J}_{\text{cond}} \) is the contribution of conducting charge carriers in the matter, the bounded current density is given by

\[ \mathbf{J}_{\text{bound}} = \mathbf{J}_{\text{pol}} + \mathbf{J}_{\text{mag}}, \quad (2.9) \]

the term \( \mathbf{J}_{\text{pol}} \), which is given by

\[ \mathbf{J}_{\text{pol}} = \frac{\partial \mathbf{P}}{\partial t}, \quad (2.10) \]

represents the contribution of dielectric polarization, the term \( \mathbf{J}_{\text{mag}} \), which is given by

\[ \mathbf{J}_{\text{mag}} = c \nabla \times \mathbf{M}, \quad (2.11) \]

represents the contribution of magnetization, and \( \mathbf{J}_{\text{ext}} \) represents all other contributions. In this notation the Maxwell’s equations (2.2), (2.3), (2.4) and (2.5) become

\[ \nabla \times \mathbf{H} = \frac{4\pi}{c} (\mathbf{J}_{\text{ext}} + \mathbf{J}_{\text{cond}}) + \frac{1}{c} \frac{\partial \mathbf{D}}{\partial t}, \quad (2.12) \]

\[ \nabla \times \mathbf{E} = -\frac{1}{c} \frac{\partial \mathbf{B}}{\partial t}, \quad (2.13) \]

\[ \nabla \cdot \mathbf{D} = 4\pi \rho_{\text{ext}}, \quad (2.14) \]

\[ \nabla \cdot \mathbf{B} = 0. \quad (2.15) \]

where \( \mathbf{E} \) is the electric field, which is related to the total charge density \( \rho_{\text{total}} \) by [see Eq. (2.4)]

\[ \nabla \cdot \mathbf{E} = 4\pi \rho_{\text{total}}, \quad (2.16) \]

\( \mathbf{B} \) is the magnetic induction, \( \mathbf{D} \), which is given by

\[ \mathbf{D} = \mathbf{E} + 4\pi \mathbf{P}, \quad (2.17) \]

is the electric displacement and \( \mathbf{H} \), which is given by

\[ \mathbf{H} = \mathbf{B} - 4\pi \mathbf{M}, \quad (2.18) \]
2.2 The Potential 4-vector

is the magnetic field. With the help of Eqs. (2.8), (2.10), (2.11), (2.12), (2.17) and (2.18) one finds that

\[ \nabla \times \mathbf{H} = \frac{4\pi}{c} (J_{\text{total}} - J_{\text{mag}}) + \frac{1}{c} \frac{\partial \mathbf{E}}{\partial t}, \]  

(2.19)

and

\[ \nabla \times \mathbf{B} = \frac{4\pi}{c} \mathbf{J}_{\text{total}} + \frac{1}{c} \frac{\partial \mathbf{E}}{\partial t}. \]  

(2.20)

Exercise 2.1.1. current conservation - Show that

\[ \nabla \cdot (\mathbf{J}_{\text{ext}} + \mathbf{J}_{\text{cond}}) + \frac{\partial \rho_{\text{ext}}}{\partial t} = 0. \]  

(2.21)

Solution 2.1.1. This relation, which is the in-matter version of the continuity equation (1.116), can be proven by applying \( \nabla \) on Eq. (2.12) and by employing Eq. (2.14).

2.2 The Potential 4-vector

Note that the set of Maxwell’s equations in medium contains two homogeneous Eqs. \( \nabla \times \mathbf{E} = \left(-\frac{1}{c}\right) \frac{\partial \mathbf{B}}{\partial t} \) (2.13) and \( \nabla \cdot \mathbf{B} = 0 \) (2.15), which are identical to the Maxwell’s equations in vacuum (1.118) and (1.120), respectively. In addition, the set of Maxwell’s equations in medium contains two inhomogeneous Eqs. (2.12) and (2.14). These Eqs. can be related to the corresponding Maxwell’s equations in vacuum \( \nabla \times \mathbf{B} = \left(4\pi/c\right) \mathbf{J} + \left(1/c\right) \frac{\partial \mathbf{E}}{\partial t} \) (1.117) and \( \nabla \cdot \mathbf{E} = 4\pi \rho \) (1.119) by the transformation \( \mathbf{E} \rightarrow \mathbf{D}, \mathbf{B} \rightarrow \mathbf{H} \) and \( \mathbf{J} \rightarrow \mathbf{J}_{\text{ext}} \), where \( \mathbf{J}_{\text{ext}} \) is defined by [compare with Eq. (1.111)]

\[ \mathbf{J}_{\text{ext}} = (c\rho_{\text{ext}}, \mathbf{J}_{\text{ext}} + \mathbf{J}_{\text{cond}})^T. \]  

(2.22)

The Maxwell’s equation \( \nabla \cdot \mathbf{B} = 0 \) (2.15) implies the existence of a 3-vector \( \mathbf{A} \) such that

\[ \mathbf{B} = \nabla \times \mathbf{A}. \]  

(2.23)

In terms of \( \mathbf{A} \), which is called the 3-vector potential, the Maxwell’s equation \( \nabla \times \mathbf{E} = \left(-1/c\right) \frac{\partial \mathbf{B}}{\partial t} \) (2.13) can be written as

\[ \nabla \times \left( \mathbf{E} + \frac{1}{c} \frac{\partial \mathbf{A}}{\partial t} \right) = 0, \]  

(2.24)

which implies the existence of a scalar \( \phi \) such that

\[ \mathbf{E} = -\nabla \phi - \frac{1}{c} \frac{\partial \mathbf{A}}{\partial t}. \]  

(2.25)
Chapter 2. Maxwell’s Equations in Matter

When the fields $\mathbf{E}$ and $\mathbf{B}$ are expressed in terms of $\phi$ and $\mathbf{A}$, the Maxwell’s equations (2.13) and (2.15) are guaranteed to be satisfied provided that $\phi$ and $\mathbf{A}$ are smooth. Note that the above relation (2.25) generalizes Eq. (1.2), which is valid only in the electrostatics case.

The potential 4-vector $\mathbf{A}$ is defined by

\[ \mathbf{A} = (\phi, A_1, A_2, A_3)^T = (\phi, \mathbf{A})^T. \]  

(2.26)

The quantity $\partial \mathbf{A}$, which is given by

\[ \partial \mathbf{A} = \frac{1}{c} \frac{\partial \phi}{\partial t} + \nabla \cdot \mathbf{A}, \]  

(2.27)

is Lorentz invariant provided that $\mathbf{A}$ is Lorentz transformed according to $\mathbf{A}' = \Lambda \mathbf{A}$ [see Eq. (1.114)].

Claim. The relations (2.23) and (2.25) can be expressed as

\[ \partial^T \mathbf{A}^T \eta = (\partial^T \mathbf{A}^T \eta)^T = \hat{F}, \]  

(2.28)

where the $4 \times 4$ matrix $\hat{F}$ is given by

\[ \hat{F} = \begin{pmatrix} 0 & E_1 & E_2 & E_3 \\ -E_1 & 0 & -B_3 & B_2 \\ -E_2 & B_3 & 0 & -B_1 \\ -E_3 & -B_2 & B_1 & 0 \end{pmatrix}. \]  

(2.29)

Proof. The following holds [see Eqs. (1.14), (1.113) and (2.26)]

\[ \partial^T \mathbf{A}^T \eta = \begin{pmatrix} \frac{\partial \phi}{\partial x_1} \\ \frac{\partial \phi}{\partial x_2} \\ \frac{\partial \phi}{\partial x_3} \end{pmatrix} \begin{pmatrix} \phi - A_1 - A_2 - A_3 \end{pmatrix} = \begin{pmatrix} \frac{1}{c} \frac{\partial \phi}{\partial t} - \frac{1}{c} \frac{\partial A_1}{\partial t} - \frac{1}{c} \frac{\partial A_2}{\partial t} - \frac{1}{c} \frac{\partial A_3}{\partial t} \\ - \frac{\partial A_1}{\partial x_1} - \frac{\partial A_2}{\partial x_2} - \frac{\partial A_3}{\partial x_3} \\ - \frac{\partial A_1}{\partial x_2} - \frac{\partial A_2}{\partial x_1} - \frac{\partial A_3}{\partial x_3} \\ - \frac{\partial A_1}{\partial x_3} - \frac{\partial A_2}{\partial x_3} - \frac{\partial A_3}{\partial x_1} \end{pmatrix}, \]  

(2.30)

thus

\[ \partial^T \mathbf{A}^T \eta - (\partial^T \mathbf{A}^T \eta)^T = \begin{pmatrix} \frac{\partial \phi}{\partial x_1} + \frac{1}{c} \frac{\partial A_1}{\partial t} & \frac{\partial \phi}{\partial x_2} + \frac{1}{c} \frac{\partial A_2}{\partial t} & \frac{\partial \phi}{\partial x_3} + \frac{1}{c} \frac{\partial A_3}{\partial t} \\ \frac{\partial A_1}{\partial x_1} & \frac{\partial A_2}{\partial x_2} & \frac{\partial A_3}{\partial x_3} \\ \frac{\partial A_1}{\partial x_2} & \frac{\partial A_2}{\partial x_1} & 0 \\ \frac{\partial A_1}{\partial x_3} & \frac{\partial A_2}{\partial x_3} & \frac{\partial A_3}{\partial x_2} \end{pmatrix}, \]  

(2.31)

in agreement with Eq. (2.28).

The above result (2.28), which expresses $\hat{F}$ as a 4-curl acting on the potential 4-vector $\mathbf{A}$, implies the following:
Claim. The field matrix $\hat{F}$ is Lorentz transformed according to
\[
\hat{F} = \Lambda^T \hat{F}' \Lambda,
\] (2.32)
provided that $A$ is Lorentz transformed according to $A' = \Lambda A$.

Proof. The assumption $A' = \Lambda A$ leads to [recall Eq. (1.114), which reads $\partial' = \partial \Lambda^{-1}$ and Eq. (1.21), which reads $\Lambda^{-1} = \eta \Lambda^T \eta$]
\[
\partial^T A^T \eta = \Lambda^T \left( (\partial')^T (A')^T \eta \right) A,
\] (2.33)
and thus
\[
\partial^T A^T \eta - (\partial^T A^T \eta)^T = \Lambda^T \left[ \left( (\partial')^T (A')^T \eta \right) - \left( (\partial')^T (A')^T \eta \right)^T \right] \Lambda,
\] (2.34)
in agreement with Eq. (2.32).

Exercise 2.2.1. Let $u$ be the relative velocity of frame $S'$ with respect to frame $S$. Show using Eq. (2.32) that the fields $E$ and $B$ are transformed according to
\[
E = E'_{||} + \gamma (E'_{\perp} - \beta \times B'_{\perp}) ,
\] (2.35)
\[
B = B'_{||} + \gamma (B'_{\perp} + \beta \times E'_{\perp}) ,
\] (2.36)
where $V_{||}$ ($V_{\perp}$) denotes the component of a 3-vector $V$ parallel (perpendicular) to $u$ and where $\beta = u/c$.

Solution 2.2.1. When the coordinates of frame $S$ are chosen such that the velocity $u$ is pointing in the $x_1$ direction $\Lambda$ becomes [see Eq. (1.26)]
\[
\Lambda = \begin{pmatrix}
\gamma & -\gamma \beta & 0 & 0 \\
-\gamma \beta & \gamma & 0 & 0 \\
0 & 0 & 1 & 0 \\
0 & 0 & 0 & 1
\end{pmatrix},
\] (2.37)
where $\beta = u/c$ and $\gamma = 1/\sqrt{1 - \beta^2}$. The transformation (2.32) yields
\[
E_1 = E'_1 ,
\] (2.38)
\[
E_2 = \gamma (E'_2 + \beta B'_3) ,
\] (2.39)
\[
E_3 = \gamma (E'_3 - \beta B'_2) ,
\] (2.40)
\[
B_1 = B'_1 ,
\] (2.41)
\[
B_2 = \gamma (B'_2 - \beta E'_3) ,
\] (2.42)
\[
B_3 = \gamma (B'_3 + \beta E'_2) .
\] (2.43)
In a vectorial form the above can be rewritten as Eqs. (2.35) and (2.36).

Note that for the case where $B'$ vanishes Eqs. (2.35) and (2.36) coincide with Eqs. (1.107) and (1.108).
2.3 Maxwell’s Equations and Lorentz Invariance

Motivated by the above results, it is henceforth assumed that the potential 4-vector $A$ is Lorentz transformed according to

$$A' = \Lambda A.$$  \hfill (2.44)

**Claim.** The inhomogeneous Maxwell’s equations (2.12) and (2.14) can be written as [see Eq. (1.113)]

$$\partial \eta \hat{G}\eta = (4\pi/c) J^T_{\text{ext}},$$  \hfill (2.45)

where the field matrix $\hat{G}$ is given by [compare with Eq. (2.29)]

$$\hat{G} = \begin{pmatrix}
0 & D_1 & D_2 & D_3 \\
-D_1 & 0 & -H_3 & H_2 \\
-D_2 & H_3 & 0 & -H_1 \\
-D_3 & -H_2 & H_1 & 0
\end{pmatrix}. \hfill (2.46)$$

**Proof.** The following holds [see Eq. (1.14)]

$$\eta \hat{G}\eta = \begin{pmatrix}
0 & -D_1 & -D_2 & -D_3 \\
-D_1 & 0 & -H_3 & H_2 \\
-D_2 & H_3 & 0 & -H_1 \\
-D_3 & -H_2 & H_1 & 0
\end{pmatrix}, \hfill (2.47)$$

and thus [see Eq. (1.113)]

$$\partial \eta \hat{G}\eta = \left( \nabla \cdot \mathbf{D}, -\frac{1}{c} \frac{\partial \mathbf{D}}{\partial t} + \nabla \times \mathbf{H} \right),$$  \hfill (2.48)

in agreement with Eqs. (2.12) and (2.14) [see Eq. (2.22)].

**Claim.** The relation (2.45) is Lorentz invariant provided that the field matrix $\hat{G}$ is transformed according to [compare with Eq. (2.32)]

$$\hat{G}' = \Lambda^T \hat{G}' \Lambda.$$  \hfill (2.49)

**Proof.** With the help of Eq. (1.21), which reads $\Lambda^{-1} = \eta \Lambda^T \eta$, Eq. (1.112), which for the case of $J_{\text{ext}}$ becomes $J'_{\text{ext}} = \Lambda J_{\text{ext}}$, Eq. (1.114), which reads $\partial' = \partial \Lambda^{-1}$, and Eq. (2.49), which reads $\hat{G}' = \Lambda^T \hat{G}' \Lambda$, one finds that

$$\partial \eta \hat{G}\eta = \partial \eta \Lambda^T \hat{G}' \Lambda \eta$$

$$= \partial' \Lambda \eta \Lambda^T \eta \hat{G}' \Lambda \eta$$

$$= \partial' \eta \hat{G}' \Lambda \eta$$

$$= \partial' \eta \hat{G}' \left( \eta \Lambda^T \eta \right)^T$$

$$= \partial' \eta \hat{G}' \left( \Lambda^{-1} \right)^T,$$

$$\partial \eta \hat{G}\eta = \partial' \eta \hat{G}' \eta,$$  \hfill (2.50)
and
\[ J'_{\text{ext}}^T = (A^{-1} J'_\text{ext})^T = J_{\text{ext}}^T (A^{-1})^T, \]  (2.51)
and thus
\[ \partial' \eta' \vec{G}' \eta = \frac{4\pi}{c} J_{\text{ext}}^T, \]  (2.52)
i.e. the relation (2.45) is Lorentz invariant.

### 2.4 Gauge Transformation

The relation between the fields \( E \) and \( B \) and the scalar \( \phi \) and the 3-vector \( A \) potentials is given by Eqs. (2.23) and (2.25). For given fields \( E \) and \( B \), however, the potentials \( \phi \) and \( A \) are not uniquely determined by Eqs. (2.23) and (2.25), as can be demonstrated by the following transformation
\[ A \rightarrow A' = A + (\partial \psi \eta)^T, \]  (2.53)
or [see Eqs. (1.113) and (2.26)]
\[ \phi \rightarrow \phi' = \phi + \frac{1}{c} \frac{\partial \psi}{\partial t}, \]  (2.54)
\[ A \rightarrow A' = A - \nabla \psi, \]  (2.55)
where \( \psi (t, r) \) is an arbitrary smooth scalar. As can be verified by substituting into Eq. (2.28), or by substituting into Eqs. (2.23) and (2.25), the transformation given by Eq. (2.53) [or Eqs. (2.54) and (2.55)], which is called gauge transformation, keeps \( E \) and \( B \) unchanged.

### 2.5 The Lorenz and Coulomb Gauge Transformations in Vacuum

In this section the Lorenz and Coulomb gauge transformations in vacuum are discussed. The generalized Lorenz gauge will be presented in the following chapter.

**Exercise 2.5.1.** Express the Maxwell’s equations in vacuum (1.117) and (1.119) in terms of the potentials \( \phi \) and \( A \).

**Solution 2.5.1.** Substituting Eqs. (2.23) and (2.25) into the Maxwell’s equations in vacuum (1.117) and (1.119) leads with the help of the vector identity
\[ \nabla \times (\nabla \times A) = \nabla (\nabla \cdot A) - \nabla^2 A, \]  (2.56)
to

$$\Box^2 A - \nabla (\partial A) = -\frac{4\pi}{c} J, \quad (2.57)$$

$$\Box^2 \phi + \frac{1}{c^2} \frac{\partial}{\partial t} (\partial A) = -4\pi \rho, \quad (2.58)$$

where $\Box^2$, which is defined by

$$\Box^2 = -\frac{1}{c^2} \frac{\partial^2}{\partial t^2} + \nabla^2, \quad (2.59)$$

is the D’Alembertian operator, and where $\partial A$ is given by [see Eq. (2.27)].

$$\partial A = \frac{1}{c} \frac{\partial \phi}{\partial t} + \nabla \cdot A. \quad (2.60)$$

### 2.5.1 Lorenz Gauge

The choice, for which

$$0 = \frac{1}{c} \frac{\partial \phi}{\partial t} + \nabla \cdot A, \quad (2.61)$$

is called the Lorenz gauge. As can be seen from Eq. (2.27), the right hand side of Eq. (2.61), which is called the Lorenz condition, is the Lorentz invariant scalar $\partial A$. For this case the Maxwell’s equations in vacuum (2.57) and (2.58) become

$$\Box^2 A = -\frac{4\pi}{c} J, \quad (2.62)$$

$$\Box^2 \phi = -4\pi \rho. \quad (2.63)$$

In electrostatics the Poisson’s equation (1.5), which is given by $\nabla^2 \phi = -4\pi \rho$, relates the 0’th component of the potential 4-vector $A = (\phi, A_1, A_2, A_3)^T$ with the 0’th component of the current 4-vector $J = (\rho, J_1, J_2, J_3)^T$. The Poisson’s equation is clearly not Lorentz invariant. The above result (2.63) generalizes it into a Lorentz invariant form.

### 2.5.2 Coulomb Gauge

Another popular choice is the Coulomb gauge, for which the following holds

$$0 = \nabla \cdot A. \quad (2.64)$$

For this case the Maxwell’s equations in vacuum (2.57) and (2.58) become

$$\left( -\frac{1}{c^2} \frac{\partial^2}{\partial t^2} + \nabla^2 \right) A - \frac{1}{c} \frac{\partial}{\partial t} \nabla \phi = -\frac{4\pi}{c} J, \quad (2.65)$$

$$\nabla^2 \phi = -4\pi \rho. \quad (2.66)$$
2.6 Integral Representation and Boundary Conditions

By applying the Stoke’s theorem, which relates a surface integral over $S$ to the closed curve integral over the boundary $C$ of the surface $S$ by

$$
\int_S (\nabla \times \mathbf{V}) \cdot d\mathbf{s} = \oint_C \mathbf{V} \cdot d\mathbf{l},
$$

(2.67)
on Eqs. (2.12) and (2.13) and the divergence theorem, which relates a volume integral over the volume $V$ to the surface integral over the boundary $S$ of the volume $V$ by

$$
\int_V (\nabla \cdot \mathbf{V}) \, dv = \int_S \mathbf{V} \cdot d\mathbf{s},
$$

(2.68)
on Eqs. (2.14) and (2.15) one obtains integral representation of the Maxwell’s equation

$$
\oint_C \mathbf{H} \cdot d\mathbf{l} = \frac{4\pi}{c} \int_S (\mathbf{J}_{\text{ext}} + \mathbf{J}_{\text{cond}}) \cdot d\mathbf{s} + \frac{1}{c} \frac{\partial}{\partial t} \int_S \mathbf{D} \cdot d\mathbf{s},
$$

(2.69)
$$
\oint_C \mathbf{E} \cdot d\mathbf{l} = -\frac{1}{c} \frac{\partial}{\partial t} \int_S \mathbf{B} \cdot d\mathbf{s},
$$

(2.70)
$$
\int_S \mathbf{D} \cdot d\mathbf{s} = 4\pi \int_V \rho_{\text{ext}} \, dv,
$$

(2.71)
$$
\int_S \mathbf{B} \cdot d\mathbf{s} = 0.
$$

(2.72)

Consider an interface between two materials. Let $\rho_s$ be the areal charge density (associated with $\rho_{\text{ext}}$) and let $\mathbf{J}_s$ be the surface current density (associated with $\mathbf{J}_{\text{ext}} + \mathbf{J}_{\text{cond}}$) on the boundary surface. For example, for the case where the boundary is taken to be the $z = 0$ plane one has

$$
\rho_{\text{ext}} = \rho_s \delta(z),
$$

(2.73)
$$
\mathbf{J}_{\text{ext}} + \mathbf{J}_{\text{cond}} = \mathbf{J}_s \delta(z).
$$

(2.74)

Let $\hat{n}$ be a unit vector normal to the interface between the two material, which are labelled as 1 and 2. In general, with the help of the vector identity

$$
\mathbf{A} \times (\mathbf{B} \times \mathbf{C}) = (\mathbf{A} \cdot \mathbf{C}) \mathbf{B} - (\mathbf{A} \cdot \mathbf{B}) \mathbf{C},
$$

(2.75)

one finds that any given vector $\mathbf{V}$ can be decomposed into parallel to $\hat{n}$ component $\mathbf{V}_n$ and a perpendicular component $\mathbf{V}_t$ according to

$$
\mathbf{V} = \mathbf{V}_n + \mathbf{V}_t,
$$

(2.76)

where
Chapter 2. Maxwell’s Equations in Matter

\[ V_n = \hat{n} (\hat{n} \cdot \mathbf{E}) , \quad (2.77) \]
\[ V_t = \hat{n} \times (\mathbf{V} \times \hat{n}) . \quad (2.78) \]

With the help of the integral representation of the Maxwell’s equation (2.69), (2.70), (2.71) and (2.72) one finds that (it is assumed that both \( \mathbf{D} \) and \( \mathbf{B} \) remain finite along the interface)

\[ \hat{n} \times (\mathbf{H}_2 - \mathbf{H}_1) = \frac{4\pi}{c} J_s , \quad (2.79) \]
\[ \hat{n} \times (\mathbf{E}_2 - \mathbf{E}_1) = 0 , \quad (2.80) \]
\[ \hat{n} \cdot (\mathbf{D}_2 - \mathbf{D}_1) = 4\pi \rho_s , \quad (2.81) \]
\[ \hat{n} \cdot (\mathbf{B}_2 - \mathbf{B}_1) = 0 . \quad (2.82) \]

**Exercise 2.6.1.** Find the boundary conditions on the surface of a perfect conductor.

**Solution 2.6.1.** Inside a perfect conductor (\( \sigma \rightarrow \infty \)) all fields vanish, and consequently the boundary conditions (2.79), (2.80), (2.81) and (2.82) become

\[ \hat{n} \times \mathbf{H} = \frac{4\pi}{c} J_s , \quad (2.83) \]
\[ \hat{n} \times \mathbf{E} = 0 , \quad (2.84) \]
\[ \hat{n} \cdot \mathbf{D} = 4\pi \rho_s , \quad (2.85) \]
\[ \hat{n} \cdot \mathbf{B} = 0 . \quad (2.86) \]

2.7 Isotropic and Linear Medium

For an isotropic and linear medium the following relations hold (in the rest frame of the medium)

\[ \mathbf{D} = \epsilon \mathbf{E} , \quad (2.87) \]
\[ \mathbf{P} = \chi_e \epsilon \mathbf{E} , \quad (2.88) \]
\[ \mathbf{B} = \mu \mathbf{H} , \quad (2.89) \]
\[ \mathbf{M} = \chi_m \mathbf{H} , \quad (2.90) \]

where \( \epsilon \) is the relative permittivity, \( \chi_e \) is the electric susceptibility, \( \mu \) is the relative permeability and \( \chi_m \) is the magnetic susceptibility, where [see Eqs. (2.17) and (2.18)]

\[ \epsilon = 1 + 4\pi \chi_e , \quad (2.91) \]
\[ \mu = 1 + 4\pi \chi_m . \quad (2.92) \]

The contribution of conducting charge carries to the current density \( \mathbf{J}_{\text{cond}} \) is related to \( \mathbf{E} \) by

\[ \mathbf{J}_{\text{cond}} = \sigma \mathbf{E} , \quad (2.93) \]

where \( \sigma \) is the conductivity.

Eyal Buks  Wave Phenomena - Lecture Notes  44
2.7. Isotropic and Linear Medium

Exercise 2.7.1. energy conservation - Show that for the case where \( J_{\text{ext}} = 0 \) the following holds

\[
\oint_S S \cdot ds + \int_V \sigma E^2 dv + \frac{\partial}{\partial t} \int_V u dv = 0 ,
\]

(2.94)

where \( S \) is the Poynting vector, which is given by

\[
S = \frac{c}{4\pi} E \times H ,
\]

(2.95)

and where \( u \) is the electromagnetic energy density, which is given by

\[
u = \frac{\varepsilon E^2 + \mu H^2}{8\pi} .
\]

(2.96)

Solution 2.7.1. By multiplying Eq. (2.12) by \( E \), multiplying Eq. (2.13) by \( H \), subtracting and employing the vector identity (1.131), which is given by

\[
\nabla \cdot (A \times B) = B \cdot (\nabla \times A) - A \cdot (\nabla \times B) ,
\]

(2.97)

one obtains

\[
\nabla \cdot S + E \cdot J_{\text{cond}} + \frac{1}{4\pi} \left( E \cdot \frac{\partial D}{\partial t} + H \cdot \frac{\partial B}{\partial t} \right) = 0 ,
\]

(2.98)

or in terms of \( u \) [see Eq. (2.93)]

\[
\nabla \cdot S + \sigma E^2 + \frac{\partial u}{\partial t} = 0 .
\]

(2.99)

Applying the divergence theorem (2.68) leads to Eq. (2.94).

Exercise 2.7.2. Maxwell-Minkowski equations - Consider an isotropic and linear medium. Show that in an inertial frame moving at velocity \( u \) with respect to the medium the following holds

\[
D + \beta \times H = \varepsilon (E + \beta \times B) ,
\]

(2.100)

\[
B - \beta \times E = \mu (H - \beta \times D) ,
\]

(2.101)

where \( \beta = u/c \).

Solution 2.7.2. Let \( S' \) be the rest frame of the medium, in which the following constitutive relations hold [see Eqs. (2.87) and (2.89)]

\[
D' = \varepsilon E' ,
\]

(2.102)

\[
B' = \mu H' .
\]

(2.103)

The following holds [see Eqs. (2.35) and (2.36)]

\[
E' = E_\parallel + \gamma (E_\perp + \beta \times B_\perp) ,
\]

(2.104)

\[
B' = B_\parallel + \gamma (B_\perp - \beta \times E_\perp) ,
\]

(2.105)
Chapter 2. Maxwell’s Equations in Matter

and [see Eq. (2.49)]

\[
D' = D + \gamma (D_{\perp} + \beta \times H_{\perp}) , \\
H' = H + \gamma (H_{\perp} - \beta \times D_{\perp}) ,
\]

(2.106) (2.107)

where \( V_{\parallel} (V_{\perp}) \) denotes the component of a 3-vector \( V \) parallel (perpendicular) to \( u \) and where \( \beta = u / c \), and thus

\[
D_{\parallel} + \gamma (D_{\perp} + \beta \times H_{\perp}) = \epsilon \left[ E_{\parallel} + \gamma (E_{\perp} + \beta \times B_{\perp}) \right] , \\
B_{\parallel} + \gamma (B_{\perp} - \beta \times E_{\perp}) = \mu \left[ H_{\parallel} + \gamma (H_{\perp} - \beta \times D_{\perp}) \right] .
\]

(2.108) (2.109)

By dividing the perpendicular components of both sides of both Eqs. (2.108) and (2.109) by \( \gamma \) one obtains

\[
D_{\parallel} + D_{\perp} + \beta \times H_{\perp} = \epsilon \left[ E_{\parallel} + E_{\perp} + \beta \times B_{\perp} \right] , \\
B_{\parallel} + B_{\perp} - \beta \times E_{\perp} = \mu \left[ H_{\parallel} + H_{\perp} - \beta \times D_{\perp} \right] ,
\]

(2.110) (2.111)

and thus Eqs. (2.100) and (2.101) hold (note that \( V_{\parallel} + V_{\perp} = V \) and \( \beta \times V_{\perp} = \beta \times V \)).

Claim. The inhomogeneous Maxwell’s equations (2.12) and (2.14) in an inertial frame moving at velocity \( u = (u_1, u_2, u_3) \) with respect to an isotropic and linear medium can be written as [compare with Eq. (2.45)]

\[
\partial g \hat{F} = \frac{4\pi}{c} J_{\text{ext}}^T ,
\]

(2.112)

where \( \hat{F} \) is given by Eq. (2.29), which reads

\[
\hat{F} = \begin{pmatrix}
0 & E_1 & E_2 & E_3 \\
-E_1 & 0 & -B_3 & B_2 \\
-E_2 & B_3 & 0 & -B_1 \\
-E_3 & -B_2 & B_1 & 0
\end{pmatrix},
\]

(2.113)

the effective metric \( g \) is given by

\[
g = \frac{1}{\sqrt{\rho}} \left( \eta + \frac{\xi}{c^2} U U^T \right) ,
\]

(2.114)

\( \eta \) is the Minkowski metric (1.14), the velocity 4-vector \( U \) is defined by [compare with Eq. (1.63)]

\[
U = \frac{dX}{d\tau} = \gamma \left( c, u_1, u_2, u_3 \right)^T .
\]

(2.115)

where \( \gamma = 1/\sqrt{1 - (w^2/c^2)} \), the parameter \( \xi \) is given by

\[
\xi = \epsilon \mu - 1 ,
\]

(2.116)

\( \epsilon \) is the relative permittivity, \( \mu \) is the relative permeability and the 4-vector \( J_{\text{ext}} = (c \rho_{\text{ext}}, J_{\text{ext}} + J_{\text{cond}})^T \) is defined by Eq. (2.22).
Proof. Let $S'$ be the rest frame of the medium. The constitutive relations $D' = \varepsilon E'$ (2.87) and $B' = \mu H'$ (2.89) can be expressed as [see Eqs. (2.29) and (2.46)]

$$\hat{G}' = \zeta \hat{F}' \zeta ,$$  
(2.117)

where

$$\hat{F}' = \begin{pmatrix} 0 & E'_1 & E'_2 & E'_3 \\ -E'_1 & 0 & -B'_3 & B'_2 \\ -E'_2 & B'_3 & 0 & -B'_1 \\ -E'_3 & -B'_2 & B'_1 & 0 \end{pmatrix} ,$$  
(2.118)

and

$$\hat{G}' = \begin{pmatrix} 0 & D'_1 & D'_2 & D'_3 \\ -D'_1 & 0 & -H'_3 & H'_2 \\ -D'_2 & H'_3 & 0 & -H'_1 \\ -D'_3 & -H'_2 & H'_1 & 0 \end{pmatrix} ,$$  
(2.119)

and where the matrix $\zeta$ is given by

$$\zeta = \frac{1}{\sqrt{\mu}} \begin{pmatrix} 1 + \xi & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} ,$$  
(2.120)

where $\xi = \epsilon \mu - 1$. Inverting the Lorentz transformations (2.32), which is given by $\hat{F} = A^T \hat{F}' A$, and (2.49) yields

$$\hat{F}' = (A^{-1})^T \hat{F} A^{-1} ,$$  
(2.121)

$$\hat{G}' = (A^{-1})^T \hat{G} A^{-1} ,$$  
(2.122)

and thus Eq. (2.117) can be rewritten as

$$\hat{G} = A^T \zeta (A^{-1})^T \hat{F} A^{-1} \zeta A .$$  
(2.123)

The above result (2.123) allows writing the Maxwell’s equation $\partial \eta \hat{G} \eta = (4\pi/c) J_{\text{ext}}^T$ (2.45) as

$$\partial g^T \hat{F} g = \frac{4\pi}{c} J_{\text{ext}}^T ,$$  
(2.124)

where

$$g = A^{-1} \zeta A \eta .$$  
(2.125)

When the Lorentz transformation $A$ is taken to be given by the matrix $B (\beta)$ [see Eq. (1.41)] one finds that

$$A^{-1} \zeta A \eta = \frac{1}{\sqrt{\mu}} \left( \eta + \frac{\xi}{c^2} U U^T \right) ,$$  
(2.126)
in agreement with Eq. (2.114). Note that \( g^T = g \). In a matrix form the metric \( g \) (2.114) is given by

\[
g = \frac{1}{\sqrt{\mu}} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} + \frac{\xi}{1 - \beta^2} \begin{pmatrix} 1 & \beta_1 & \beta_2 & \beta_3 \\ \beta_1 & \beta_1^2 & \beta_1 \beta_2 & \beta_1 \beta_3 \\ \beta_2 & \beta_2 \beta_1 & \beta_2^2 & \beta_2 \beta_3 \\ \beta_3 & \beta_3 \beta_1 & \beta_3 \beta_2 & \beta_3^2 \end{pmatrix},
\]

where \( \beta \) is related to the velocity 3-vector \( \mathbf{u} \) by \( \beta = \mathbf{u}/c \) [see Eq. (2.115)].

With the help of the relation \( \partial^T A^T \eta = (\partial^T A^T \eta)^T = \hat{F} \) [see Eq. (2.28)] Eq. (2.112) can be expressed in terms of the potential 4-vector \( A \) as

\[
\partial g \left( \partial^T A^T \eta - (\partial^T A^T \eta)^T \right) g = \frac{4\pi}{c} J_{ext}^T.
\]

### 2.8 Harmonic Time Dependency

Consider a monochromatic solution of the Maxwell’s equations, for which all fields and sources oscillate in time at angular frequency \( \omega \). It is convenient to employ complex notation, in which all fields and sources are expressed as

\[
\psi(r, t) = \text{real} \left[ \psi(r) e^{-i\omega t} \right].
\]

Note that in order to avoid cumbersome notation, the same letter \( \psi \) denotes the \( r \) and \( t \) dependent amplitude \( \psi(r, t) \) and the \( r \) only dependent amplitude \( \psi(r) \), which is commonly called a phasor.

By substituting into the Maxwell’s equations (2.12), (2.13), (2.14) and (2.15) one finds for the case of isotropic and linear response that

\[
\nabla \times \mathbf{H} = \frac{4\pi J_{ext}}{c} - \frac{i\omega}{c} \epsilon_{eff} \mathbf{E},
\]

\[
\nabla \times \mathbf{E} = \frac{i\omega}{c} \mu \mathbf{H},
\]

\[
\nabla \cdot (\epsilon \mathbf{E}) = 4\pi \rho_{ext},
\]

\[
\nabla \cdot (\mu \mathbf{H}) = 0,
\]

where the effective relative dielectric coefficient \( \epsilon_{eff} \) is given by

\[
\epsilon_{eff} = \epsilon + \frac{4\pi \sigma}{\omega}.
\]

**Exercise 2.8.1.** Consider two vectors \( \mathbf{V}_a \) and \( \mathbf{V}_b \) having harmonic time dependency

\[
\mathbf{V}_a(r, t) = \text{real} \left[ \mathbf{V}_a(r) e^{-i\omega t} \right],
\]

\[
\mathbf{V}_b(r, t) = \text{real} \left[ \mathbf{V}_b(r) e^{-i\omega t} \right].
\]

Calculate the time averaged of \( \mathbf{V}_a(r, t) \cdot \mathbf{V}_b(r, t) \).
Solution 2.8.1. The symbol \( \langle \rangle \) is employed below to label time averaging [it should not be confused with the spacial averaging that has been defined by Eq. (2.1), even though the same symbol is employed]. The following holds

\[
\langle V_a(\mathbf{r}, t) \cdot V_b(\mathbf{r}, t) \rangle = \sum_{n=1}^{3} \langle \text{real} \left[ V_{a,n} e^{-i\omega t} \right] \cdot \text{real} \left[ V_{b,n} e^{-i\omega t} \right] \rangle
\]

\[
= \sum_{n=1}^{3} \left\langle \frac{V_{a,n} e^{-i\omega t} + V_{a,n}^* e^{i\omega t}}{2} \cdot \frac{V_{b,n} e^{-i\omega t} + V_{b,n}^* e^{i\omega t}}{2} \right\rangle
\]

\[
= \frac{1}{2} \sum_{n=1}^{3} \text{real} \left( V_{a,n} V_{b,n}^* \right), \quad (2.137)
\]

thus

\[
\langle V_a(\mathbf{r}, t) \cdot V_b(\mathbf{r}, t) \rangle = \frac{1}{2} \text{real} \left( V_a \cdot V_b^* \right) \quad (2.138)
\]

2.9 Inhomogeneous Medium Free of Sources

Consider the case of electromagnetic fields in an inhomogeneous medium free of sources (i.e. \( \rho_{\text{ext}} = 0 \) and \( \mathbf{J}_{\text{ext}} = 0 \)), which is assumed to be isotropic, linear and stationary. In addition, the conductivity \( \sigma \) is assumed to vanish and \( \epsilon = \epsilon(\mathbf{r}) \) and \( \mu = \mu(\mathbf{r}) \) are taken to be time independent scalars. For that case Eqs. (2.130), (2.131), (2.132) and (2.133) become

\[
\nabla \times \mathbf{H} = -ik_0 \epsilon \mathbf{E}, \quad (2.140)
\]

\[
\nabla \times \mathbf{E} = i k_0 \mu \mathbf{H}, \quad (2.141)
\]

\[
\nabla \cdot (\epsilon \mathbf{E}) = 0, \quad (2.142)
\]

\[
\nabla \cdot (\mu \mathbf{H}) = 0, \quad (2.143)
\]

where

\[
k_0 = \frac{\omega}{c}. \quad (2.144)
\]

Note that Eqs. (2.142) and (2.143) result from Eqs. (2.140) and (2.141) by the vector identity

\[
\nabla \cdot (\nabla \times \mathbf{A}) = 0. \quad (2.145)
\]

Exercise 2.9.1. Show that

\[
\nabla^2 \mathbf{E} + n^2 k_0^2 \mathbf{E} + (\nabla \log \mu) \times (\nabla \times \mathbf{E}) + \nabla (\mathbf{E} \cdot \nabla \log \epsilon) = 0, \quad (2.146)
\]

\[
\nabla^2 \mathbf{H} + n^2 k_0^2 \mathbf{H} + (\nabla \log \epsilon) \times (\nabla \times \mathbf{H}) + \nabla (\mathbf{H} \cdot \nabla \log \mu) = 0, \quad (2.147)
\]

where \( n \), which is given by
\[ n = \sqrt{\frac{\epsilon}{\mu}}, \quad (2.148) \]

is the so-called refraction index.

**Solution 2.9.1.** Applying the operator \( \nabla \times \) to Eq. (2.140) and using the vector identity (2.56), which is given by

\[ \nabla \times (\nabla \times A) = \nabla (\nabla \cdot A) - \nabla^2 A, \quad (2.149) \]

one finds that

\[ \nabla (\nabla \cdot H) - \nabla^2 H = - \frac{i\omega}{c} \nabla \times (\epsilon E). \quad (2.150) \]

Using the vectors identities

\[ \nabla \cdot (f A) = f \nabla \cdot A + A \cdot \nabla f, \quad (2.151) \]

and

\[ \nabla \times (f A) = f \nabla \times A + (\nabla f) \times A, \quad (2.152) \]

together with Eqs. (2.141) and (2.143) leads to Eq. (2.147). Equation (2.146) is obtained in a similar way.

### 2.10 The Scalar Approximation

In the scalar approximation the third and forth terms on the right hand sides of Eqs. (2.146) and (2.147) are disregarded. These terms give rise to coupling between the components of \( E \) and \( H \). After the removal of these terms Eqs. (2.146) and (2.147) imply that all three components of \( E \) and \( H \) satisfy the so-called Helmholtz equation, which is given by

\[ (\nabla^2 + n^2 k_0^2) \psi = 0. \quad (2.153) \]

Note that for the case of a homogeneous medium, in which both \( \epsilon \) and \( \mu \) are constants, the above statement [i.e. all three components of \( E \) and \( H \) satisfy the Helmholtz equation (2.153)] becomes exact.

### 2.11 Polarization

Consider an electric field, which is denoted in this section by \( E_t \), having harmonic time dependency

\[ E_t = \text{real} \left[ \exp (-i\omega t) E \right], \quad (2.154) \]
where the complex phasor vector $\mathbf{E}$ is taken to be given by

$$
\mathbf{E} = a + i b ,
$$

(2.155)

where $a$ and $b$ are real vectors (constants for a given spatial position). The real time dependent field is given by

$$
\mathbf{E}_t = \frac{1}{2} [\exp (-i\omega t) \mathbf{E} + \exp (i\omega t) \mathbf{E}^*] \\
= \frac{1}{2} [\exp (-i\omega t) (a + i b) + \exp (i\omega t) (a - i b)] \\
= a \cos (\omega t) + b \sin (\omega t) .
$$

(2.156)

Clearly $\mathbf{E}_t (t)$ is a close, planar and periodic curve, lying in the plane perpendicular to $a \times b$.

**Claim.** The curve $\mathbf{E}_t (t)$ is an ellipse.

**Proof.** To show that $\mathbf{E}_t (t) = (E_x, E_y, E_z)$ is an ellipse one needs to show that it is a conic section (see Fig. 2.1), namely, the components $E_x, E_y, E_z$ satisfy a 2nd order equation of the type

$$
0 \leq \sum_{0 \leq n_x+n_y+n_z \leq 2} A_{n_x,n_y,n_z} E_x^{n_x} E_y^{n_y} E_z^{n_z} = 0 ,
$$

(2.157)

where all coefficients $A_{n_x,n_y,n_z}$ are time independent real constants. The following holds [see Eq. (2.156)]

$$
E_t^2 (t) = a_i^2 \cos^2 (\omega t) + b_i^2 \sin^2 (\omega t) + a_i b_i \sin (2\omega t) ,
$$

(2.158)

or

$$
E_t^2 (t) - \frac{1}{2} (a_i^2 + b_i^2) = \frac{1}{2} (a_i^2 - b_i^2) \cos (2\omega t) + a_i b_i \sin (2\omega t) ,
$$

(2.159)

where $i = 1, 2, 3$. In a matrix form

$$
M \begin{pmatrix}
\cos (2\omega t) \\
\sin (2\omega t) \\
-1
\end{pmatrix} = 0 ,
$$

(2.160)

where

$$
M = \begin{pmatrix}
\frac{a_x^2-b_x^2}{a_x^2} & \frac{a_y b_x}{a_x} & \frac{E_x^2 (t) - \frac{a_x^2+b_x^2}{2}}{a_x} \\
\frac{a_y^2-b_y^2}{a_y^2} & \frac{a_x b_y}{a_y} & \frac{E_y^2 (t) - \frac{a_y^2+b_y^2}{2}}{a_y} \\
\frac{a_z^2-b_z^2}{a_z^2} & \frac{a_y b_z}{a_z} & \frac{E_z^2 (t) - \frac{a_z^2+b_z^2}{2}}{a_z}
\end{pmatrix} .
$$

(2.161)

The condition for solution existence for the 2 unknowns $\cos (2\omega t)$ and $\sin (2\omega t)$ requires that

$$
\det M = 0 ,
$$

(2.162)

thus, $\mathbf{E}_t (t)$ is indeed conic section and thus (since it is periodic) an ellipse.
The eccentricity $e$ of an ellipse is defined by

$$e = \sqrt{1 - \frac{\left(\frac{E^2}{|E|^2}\right)_{\text{min}}}{\left(\frac{E^2}{|E|^2}\right)_{\text{max}}}},$$

(2.163)

where $\left(\frac{E^2}{|E|^2}\right)_{\text{min}}$ ($\left(\frac{E^2}{|E|^2}\right)_{\text{max}}$) is the minimum (maximum) value of $\frac{E^2}{|E|^2}$.

**Exercise 2.11.1.** Show that eccentricity is given by

$$e = \sqrt{\frac{2|\eta|}{1 + |\eta|}},$$

(2.164)

where

$$\eta = \frac{E^2}{|E|^2}.$$  

(2.165)

**Solution 2.11.1.** Taking the square of $E_t(t)$ leads to [see Eq. (2.154)]

$$E^2_t(t) = \frac{1}{4} \left[ \exp(-2i\omega t) E^2 + \exp(2i\omega t) (E^*)^2 + 2E \cdot E^* \right].$$

(2.166)

Using the notation

$$\eta = \frac{E^2}{|E|^2},$$

(2.167)

where $\eta = |\eta| e^{i\vartheta}$, and $\vartheta$ is real, one finds that

$$E^2_t(t) = \frac{|E|^2}{2} \left[ 1 + |\eta| \cos(2\omega t - \vartheta) \right].$$

(2.168)
Thus, while $|\eta|$ determines the eccentricity

$$e = \sqrt{1 - \frac{(E_t^2)^{\min}}{(E_t^2)^{\max}}} = \sqrt{\frac{2|\eta|}{1 + |\eta|}},$$

(2.169)

the phase $\vartheta$ of $\eta$ determines the phase of time oscillations.

**Exercise 2.11.2.** Show that

$$(E_t^2)^{\min,\max} = \frac{1}{2} \left( a^2 + b^2 \right) \pm \frac{1}{2} \sqrt{4 (a \cdot b)^2 + (a^2 - b^2)^2}.$$  

(2.170)

**Solution 2.11.2.** In terms of $a$ and $b$ one has [see Eq. (2.156)]

$$E_t^2(t) = a^2 \cos^2(\omega t) + b^2 \sin^2(\omega t) + a \cdot b \sin(2\omega t).$$

(2.171)

The extremum points of $E_t^2$ are found by solving

$$0 = \frac{dE_t^2}{d(\omega t)} = -a^2 \sin(2\omega t) + b^2 \sin(2\omega t) + 2a \cdot b \cos(2\omega t),$$

(2.172)

thus

$$\tan(2\omega t) = \frac{2a \cdot b}{a^2 - b^2}.$$  

(2.173)

Rewriting Eq. (2.171) as

$$E_t^2(t) = \frac{1}{2} (a^2 + b^2) + \frac{1}{2} \cos(2\omega t) (a^2 - b^2) + a \cdot b \sin(2\omega t),$$

(2.174)

and using Eq. (2.173) together with the identities

$$\sin x = \pm \frac{\tan x}{\sqrt{\tan^2 x + 1}},$$

(2.175)

and

$$\cos x = \pm \frac{1}{\sqrt{\tan^2 x + 1}},$$

(2.176)

lead to

$$(E_t^2)^{\min,\max} = \frac{1}{2} (a^2 + b^2) \pm \frac{1}{2} \sqrt{\frac{1}{\left(\frac{2a \cdot b}{a^2 - b^2}\right)^2 + 1} \left[ (a^2 - b^2) + a \cdot b \frac{4a \cdot b}{a^2 - b^2} \right]},$$

$$= \frac{1}{2} (a^2 + b^2) \pm \frac{1}{2} \sqrt{4 (a \cdot b)^2 + (a^2 - b^2)^2}.$$  

(2.177)
Chapter 2. Maxwell’s Equations in Matter

2.12 Problems

1. Calculate the electric \( \mathbf{E} \) and magnetic \( \mathbf{B} \) fields generated by a point particle of charge \( q \) moving in the \( x \) direction at a fixed velocity \( u \).

2. Consider an uncharged dielectric sphere having a homogeneous permittivity \( \varepsilon \) and radius \( R \). Calculate the electric field \( \mathbf{E} \) inside and outside the sphere given that far from the sphere \( \mathbf{E} = E_0 \hat{z} \), where \( E_0 \) is a constant.

3. The Drude model - Consider a conductor containing charge carriers having charge \( q \) and mass \( m \) in the presence of electrical field \( \mathbf{E} \) (and vanishing magnetic field). The density of charge carriers (i.e. number per unit volume) is \( n_{cc} \). Scattering is taken into account in the Drude model by adding a damping term to the classical equation of motion [see Eq. (1.1)]

\[
\frac{dp}{dt} + \frac{p}{\tau_{tr}} = qE,
\]

(2.178)

where \( p \) is the momentum per electron and where \( \tau_{tr} \) is the so-called scattering time. Calculate the frequency dependent effective dielectric coefficient \( \varepsilon_{eff}(\omega) \) of the conductor.

4. Fresnel equations - Consider a plane wave of wave vector \( \mathbf{k}_i \) striking the planar interface between two lossless materials having refractive indices \( n_l = \sqrt{\varepsilon_l \mu_l} \), where for the material hosting the incident wave \( l = 1 \) and for the other material \( l = 2 \). The plane containing \( \mathbf{k}_i \) and \( \hat{n} \), where \( \hat{n} \) is a unit vector normal to the interface, is called the plane of incidence. Let \( \theta_i \) be the angle between \( \mathbf{k}_i \) and \( \hat{n} \). Calculate the fraction of the power that is reflected from the interface for the cases of s-polarization (i.e. when the electric field of the incident wave is orthogonal to the plane of incidence) and p-polarization (i.e. when the electric field of the incident wave is in the plane of incidence).

5. Consider a layer of width \( d \) and a refractive index \( n_3 = \sqrt{\varepsilon_3 \mu_3} \) sandwiched between two semi-infinite media having refractive index \( n_1 = \sqrt{\varepsilon_1 \mu_1} \) and \( n_2 = \sqrt{\varepsilon_2 \mu_2} \), respectively. Solutions of the Maxwell’s equations for which the electromagnetic fields are confined near the central layer are called surface modes. Find an equation which is satisfied by all angular frequencies \( \omega \) corresponding to surface modes having s-polarization and an equation for the case of p-polarization.

6. The Lifshitz formula - According to the theory of quantum mechanics, the ground state energy (or zero-point energy) of an electromagnetic mode having angular frequency \( \omega \) is \( \hbar \omega/2 \), where \( \hbar \) is Planck’s h-bar constant.

a) Consider the trilayer of the previous problem and calculate the mutual force (which is commonly called the Casimir force) between layer 1 and layer 2 originating by the dependence of the zero-point energy \( u(d) \) on the distance \( d \) between the layers.
2.13 Solutions

b) Calculate the Casimir force for the case where both layers 1 and 2 have metallic dielectric coefficient given by [see Eq. (2.216)]

\[
\epsilon_{1,2}(\omega) = 1 - \frac{\omega_p^2}{\omega^2},
\]

(2.179)

where \(\omega_p\) is the plasma frequency, \(\epsilon_3 = 1\) and \(\mu_1 = \mu_2 = \mu_3 = 1\).

7. Stokes parameters - Consider a monochromatic electromagnetic plane wave propagating in vacuum along the \(z\) axis. The components of the electric field vector \(E = (E_x, E_y, E_z)\) are assumed to be given by

\[
E_x = E_{x0} \cos \left( \frac{\omega z - ct}{c} + \delta_x \right), \quad (2.180)
\]

\[
E_y = E_{y0} \cos \left( \frac{\omega z - ct}{c} + \delta_y \right), \quad (2.181)
\]

\[
E_z = 0, \quad (2.182)
\]

where \(\omega, E_{x0}, E_{y0}, \delta_x\) and \(\delta_y\) are constants. The so-called Stokes parameters are defined by

\[
S_0 = E_{x0}^2 + E_{y0}^2, \quad (2.183)
\]

\[
S_1 = E_{x0}^2 - E_{y0}^2, \quad (2.184)
\]

\[
S_2 = 2E_{x0}E_{y0} \cos (\delta_x - \delta_y), \quad (2.185)
\]

\[
S_3 = 2E_{x0}E_{y0} \sin (\delta_x - \delta_y). \quad (2.186)
\]

Note that when \(S_1 = S_2 = 0\) the polarization is circular, whereas when \(S_3 = 0\) the polarization is rectilinear. Calculate the Stokes parameters \(S'_0, S'_1, S'_2\) and \(S'_3\) as measured by an observer moving at a constant velocity given by \(c\beta\), where the vector \(\beta\) is expressed as \(\beta = \beta \hat{\beta}\), where \(\beta = |\beta|\), and the unit vector \(\hat{\beta} = (\sin \theta, 0, \cos \theta)\) is assumed to lie in the \(xz\) plane.

8. Show that the scalars \(E \cdot B, E \cdot E - B \cdot B, D \cdot H, D \cdot D - H \cdot H\) and \(E \cdot D - B \cdot H\) are all Lorentz invariant.

9. The Leinard–Wiechert potential - Show that in the Lorenz gauge in vacuum the potential 4-vector \(A\) is related to the current 4-vector \(J\) by

\[
A(t, x) = \int d^3x' \frac{J \left( t - \frac{|x-x'|}{c}, x' \right)}{c|x-x'|}. \quad (2.187)
\]

2.13 Solutions

1. In a frame commoving with the particle the electric field \(E\) is given by the Coulomb’s law

\[
E' = \frac{q \Gamma'}{|\Gamma'|^3}, \quad (2.188)
\]
Chapter 2. Maxwell’s Equations in Matter

and the magnetic field vanishes, i.e. $B' = 0$, and thus [see Eqs. (2.35) and (2.36)]

$$E = E'_\parallel + \gamma (E'_1 - \beta \times B'_1)$$

$$= \frac{q x'_1 \hat{s}_1}{(x'_1^2 + x'_2^2 + x'_3^2)^{3/2}} + \gamma \frac{q (x'_2 \hat{s}_2 + x'_3 \hat{s}_3)}{(x'_1^2 + x'_2^2 + x'_3^2)^{3/2}},$$

(2.189)

and

$$B = B'_\parallel + \gamma (B'_1 + \beta \times E'_1)$$

$$= \gamma \beta \times \frac{q (x'_2 \hat{s}_2 + x'_3 \hat{s}_3)}{(x'_1^2 + x'_2^2 + x'_3^2)^{3/2}},$$

(2.190)

where

$$\beta = \frac{u}{c} \hat{s}_1,$$

(2.191)

and where

$$\gamma = \frac{1}{\sqrt{1 - \left(\frac{u}{c}\right)^2}}.$$ (2.192)

The time $t$ and $x$ coordinates are transformed according to Eq. (1.25)

$$\begin{pmatrix} ct' \\ x'_1 \end{pmatrix} = \gamma \begin{pmatrix} 1 & -\frac{u}{c} \\ -\frac{u}{c} & 1 \end{pmatrix} \begin{pmatrix} ct \\ x_1 \end{pmatrix},$$

(2.193)

and therefore

$$E = \frac{q \gamma}{r_0} x_0,$$

(2.194)

and

$$B = \frac{q \gamma}{r_0} \beta \times x_0,$$

(2.195)

where

$$x_0 = x - ut,$$

(2.196)

and where

$$r_0 = \sqrt{\gamma^2 (x_1 - ut)^2 + x_2^2 + x_3^2}.$$ (2.197)
2. The coordinates are chosen such that the center of the sphere is located at the origin. Since the magnetic induction $B$ vanishes for this problem of electrostatics $E$ can be expressed in terms of a scalar potential $\phi$ as [see Eq. (2.13)]

$$E = -\nabla \phi.$$  \hfill (2.198)

The transformation from Cartesian to spherical coordinates is given by

$$x = r \sin \theta \cos \varphi,$$  \hfill (2.199)

$$y = r \sin \theta \sin \varphi,$$  \hfill (2.200)

$$z = r \cos \theta,$$  \hfill (2.201)

where $r \geq 0$, $0 \leq \theta \leq \pi$, and $0 \leq \varphi \leq 2\pi$, and thus in spherical coordinates one has

\[
\nabla = \hat{\mathbf{r}} \frac{\partial}{\partial r} + \hat{\mathbf{\theta}} \frac{1}{r} \frac{\partial}{\partial \theta} + \hat{\mathbf{\varphi}} \frac{1}{r \sin \theta} \frac{\partial}{\partial \varphi}.
\]  \hfill (2.202)

By symmetry, the scalar potential $\phi$ is expected to be independent on $\varphi$, and thus it can be expressed as

\[
\phi (r, \theta) = \begin{cases} 
\phi_{\text{in}} (r, \theta) & r < R \\
\phi_{\text{out}} (r, \theta) & r > R 
\end{cases}.
\]  \hfill (2.203)

The requirement that $E = E_0 \hat{z}$ far from the sphere implies that $\phi_{\text{out}} \simeq -E_0 z = -E_0 r \cos \theta$ when $r \gg R$. On the surface of the sphere the boundary condition (2.80) reads

$$\hat{n} \times \nabla \phi_{\text{out}} = \hat{n} \times \nabla \phi_{\text{in}},$$  \hfill (2.204)

and the boundary condition (2.81) reads [recall Eq. (2.87), which reads $D = \epsilon E$]

$$\hat{n} \cdot \nabla \phi_{\text{out}} = \epsilon \hat{n} \cdot \nabla \phi_{\text{in}},$$  \hfill (2.205)

where $\hat{n}$ is a unit vector normal to the surface of the sphere, thus at $r = R$ the solution is required to satisfy

$$\frac{\partial \phi_{\text{out}}}{\partial \theta} = \frac{\partial \phi_{\text{in}}}{\partial \theta},$$  \hfill (2.206)

and

$$\frac{\partial \phi_{\text{out}}}{\partial r} = \epsilon \frac{\partial \phi_{\text{in}}}{\partial r}.$$  \hfill (2.207)
Chapter 2. Maxwell’s Equations in Matter

All these requirements are satisfied by

$$\frac{\phi}{E_0} = \begin{cases} \frac{3}{2} \cos \theta r < R \\ \left( \frac{\epsilon - 1}{\epsilon + 2} \frac{R^3}{r^3} - r \right) \cos \theta r > R \end{cases} \tag{2.208}$$

and thus [see Eqs. (2.198) and (2.202)]

$$\frac{E}{E_0} = \begin{cases} \hat{r} \left( 1 + \frac{2(\epsilon - 1)}{\epsilon (2 + \epsilon)} \right) \cos \theta + \hat{\theta} \left( \frac{\epsilon - 1}{\epsilon (2 + \epsilon)} - 1 \right) \sin \theta r > R \end{cases} \tag{2.209}$$

3. In terms of the current density vector $\mathbf{J}$, which is related to $\mathbf{p}$ by the relation

$$\mathbf{p} = \frac{m}{q n_{cc}} \mathbf{J} \quad \tag{2.210}$$

Eq. (2.178) yields

$$\frac{m}{q^2 n_{cc}} \left( \frac{\partial \mathbf{J}}{\partial t} + \frac{1}{\tau_{tr}} \mathbf{J} \right) = \mathbf{E} \quad \tag{2.211}$$

When harmonic time dependency at angular frequency $\omega$ is assumed Eq. (2.211) yields

$$\frac{m}{q^2 n_{cc}} \left( -i\omega + \frac{1}{\tau_{tr}} \right) \mathbf{J} (\omega) = \mathbf{E} (\omega) \quad \tag{2.212}$$

and thus the conductivity $\sigma$ is given by [see Eq. (2.93)]

$$\sigma (\omega) = \frac{\sigma_0}{1 - i\omega \tau_{tr}} \quad \tag{2.213}$$

where

$$\sigma_0 = \frac{q^2 n_{cc} \tau_{tr}}{m} \quad \tag{2.214}$$

and therefore the effective dielectric coefficient $\epsilon_{eff}$ is given by [see Eq. (2.134)]

$$\epsilon_{eff} (\omega) = 1 + \frac{4\pi i}{\omega} \frac{\sigma_0}{1 - i\omega \tau_{tr}} \quad \tag{2.215}$$

When $\omega \tau_{tr} \gg 1$ this becomes

$$\epsilon_{eff} (\omega) = 1 - \frac{\omega_p^2}{\omega^2} \quad \tag{2.216}$$

where $\omega_p$, which is given by

$$\omega_p^2 = \frac{4\pi q^2 n_{cc}}{m} \quad \tag{2.217}$$

is the so-called plasma frequency.
4. The interface between the two materials is taken to be the \( z = 0 \) plane, where \( n = n_1 \) \((n = n_2)\) for \( z < 0 \) \((z > 0)\) and the plane of incidence is taken to be the plane spanned by \( \hat{n} = \hat{z} \) and \( \hat{x} \). Consider a solution for the electric field \( \mathbf{E} \) composed of incident, reflected and refracted plane waves having wave vectors \( \mathbf{k}_i, \mathbf{k}_r \) and \( \mathbf{k}_t \), respectively. For the case of s-polarization the solution is expressed as [see Eq. (2.153)]

\[
\mathbf{E}_s = \begin{cases} 
E_i (e^{i\mathbf{k}_i \cdot \mathbf{r}} + r_s e^{i\mathbf{k}_r \cdot \mathbf{r}}) \hat{y} & z < 0 \\
t_s E_i e^{i\mathbf{k}_r \cdot \mathbf{r}} & z > 0
\end{cases},
\]

(2.218)

and for the case of p-polarization as

\[
\mathbf{E}_p = \begin{cases} 
E_i \left( e^{i\mathbf{k}_i \cdot \mathbf{r}} \frac{\mathbf{k}_i \times \hat{y}}{k_i} + r_p e^{i\mathbf{k}_r \cdot \mathbf{r}} \frac{\mathbf{k}_r \times \hat{y}}{k_r} \right) & z < 0 \\
E_{t_p} e^{i\mathbf{k}_t \cdot \mathbf{r}} \frac{\mathbf{k}_t \times \hat{y}}{k_t} & z > 0
\end{cases},
\]

(2.219)

where \( r_s \) and \( t_s \) \((r_p \) and \( t_p)\) are the reflection and transmission amplitudes, respectively; for the case of s-polarization \( (p\text{-polarization}) \) and \( E_i \) is the amplitude of incident wave. The boundary condition (2.80) can be satisfied for every point in the plane \( z = 0 \) only when \( k_i, k_r \) and \( k_t \) have the same tangential component. This requirement is satisfied by expressing \( k_i, k_r \) and \( k_t \) in terms of the corresponding angles \( \theta_i, \theta_r \) and \( \theta_t \) as [see Eq. (2.153)]

\[
k_i = \frac{n_1 \omega}{c} (\sin \theta_i, 0, \cos \theta_i) ,
\]

(2.220)

\[
k_r = \frac{n_1 \omega}{c} (\sin \theta_r, 0, -\cos \theta_r) ,
\]

(2.221)

\[
k_t = \frac{n_2 \omega}{c} (\sin \theta_t, 0, \cos \theta_t) ,
\]

(2.222)

where \( \theta_r \) is related to \( \theta_i \) by the so-called law of reflection

\[
\theta_r = \theta_i ,
\]

(2.223)

and \( \theta_t \) is related to \( \theta_i \) by the so-called Snell’s law

\[
n_1 \sin \theta_i = n_2 \sin \theta_t .
\]

(2.224)

The magnetic field \( \mathbf{H} \) is related to \( \mathbf{E} \) by Eq. (2.141), which reads

\[
\mathbf{H} = -i \frac{c}{\omega \mu} \nabla \times \mathbf{E} ,
\]

(2.225)

thus [see Eqs. (1.95), (2.152), (2.218) and (2.219)]

\[
\mathbf{H}_s = \begin{cases} 
\frac{E_{n_1}}{\mu} \left( e^{i\mathbf{k}_i \cdot r} \frac{\mathbf{k}_i \times \hat{y}}{k_i} + r_s e^{i\mathbf{k}_r \cdot r} \frac{\mathbf{k}_r \times \hat{y}}{k_r} \right) & z < 0 \\
\frac{E_{n_1}}{\mu_2} t_s e^{i\mathbf{k}_r \cdot r} \frac{\mathbf{k}_r \times \hat{y}}{k_r} & z > 0
\end{cases},
\]

(2.226)

and
\[ H_p = \begin{cases} \frac{-E_{n1}}{\mu_1} (e^{ik_x \cdot r} + r_p e^{ik_r \cdot r}) \hat{y} z < 0 \\
\frac{-E_{n2}}{\mu_2} t_p e^{ik_r \cdot r} \hat{y} z > 0 \end{cases} \]  
(2.227)

The boundary conditions (2.79) and (2.80) yield for the case of s-polarization [see Eqs. (2.218) and (2.226)]

\[ \sqrt{\frac{\epsilon_1}{\mu_1}} (1 - r_s) \cos \theta_s = \sqrt{\frac{\epsilon_2}{\mu_2}} t_s \cos \theta_t , \]  
(2.228)

\[ 1 + r_s = t_s , \]  
(2.229)

and for the case of p-polarization [see Eqs. (2.219) and (2.227)]

\[ \sqrt{\frac{\epsilon_1}{\mu_1}} (1 + r_p) = \sqrt{\frac{\epsilon_2}{\mu_2}} t_p , \]  
(2.230)

\[ (1 - r_p) \cos \theta_t = t_p \cos \theta_t . \]  
(2.231)

The solutions are given by

\[ r_s = \frac{\sqrt{\frac{\epsilon_1}{\mu_1}} \cos \theta_i - \sqrt{\frac{\epsilon_2}{\mu_2}} \cos \theta_t}{\sqrt{\frac{\epsilon_1}{\mu_1}} \cos \theta_i + \sqrt{\frac{\epsilon_2}{\mu_2}} \cos \theta_t} , \]  
(2.232)

and

\[ r_p = \frac{\sqrt{\frac{\epsilon_2}{\mu_2}} \cos \theta_i - \sqrt{\frac{\epsilon_1}{\mu_1}} \cos \theta_t}{\sqrt{\frac{\epsilon_2}{\mu_2}} \cos \theta_i + \sqrt{\frac{\epsilon_1}{\mu_1}} \cos \theta_t} . \]  
(2.233)

Note that when total internal reflection occurs, i.e. when \( n_1 > n_2 \) and [see Eq. (2.224)]

\[ \sin \theta_i > \frac{n_2}{n_1} , \]  
(2.234)

one has \( |r_s|^2 = |r_p|^2 = 1 \) [note that Eq. (2.224) yields no real solution for \( \theta_t \) for this case]. Consider the case where \( \mu_1 = \mu_2 \). For that case \( |r_p|^2 = 0 \) when \( \theta_t = \theta_B \), where \( \theta_B \) is the so-called Brewster’s angle, which is given by [see Eqs. (2.224) and (2.233)]

\[ \theta_B = \tan^{-1} \frac{n_2}{n_1} . \]  
(2.235)

5. Consider first the case of s-polarization. For this case, the solution for the electric field \( \mathbf{E} \) for each layer is assumed to have the form [compare with Eq. (2.218)]

\[ \mathbf{E} = \hat{y} e_y (z) e^{i\kappa \cdot \rho} , \]  
(2.236)

where \( \kappa = (k_x, k_y) \) and \( \rho = (x, y) \). The Helmholtz equation (2.153) implies that
\[
\frac{e''_y}{e_y} = K^2, \quad (2.237)
\]

where
\[
K^2 = \kappa^2 - \left( \frac{m\omega}{c} \right)^2. \quad (2.238)
\]

Note that for the case of surface modes \(K^2 > 0\). The magnetic field \(H\) is given by [see Eqs. (2.141) and (2.152)]
\[
H = \frac{1}{ik_0\mu} \nabla \times E = \frac{(-e'_y, 0, 0) + i(0, 0, k_x e_y)}{ik_0\mu} e^{i\kappa \cdot \rho}, \quad (2.239)
\]

where \(k_0 = \omega/c\) [see Eq. (2.144)]. Next, consider an interface between two materials at a plane of constant \(z\). The Snell’s law (2.224) implies that the lateral wave vector \(\kappa\) obtains the same value on both sides of the interface. Thus, for this case of s-polarization the boundary condition (2.80) implies that \(e_y'\) is continuous, and the boundary condition (2.79) implies that \(\mu^{-1}e_y\) is continuous. For the trilayer, the refractive index \(n\) is taken to be given by
\[
n(z) = \begin{cases} 
\sqrt{\varepsilon_1 \mu_1} & z < 0 \\
\sqrt{\varepsilon_3 \mu_3} & 0 \leq z \leq d \\
\sqrt{\varepsilon_2 \mu_2} & z < d
\end{cases}, \quad (2.240)
\]

Consider a solution having the form [see Eq. (2.237)]
\[
e_y(z) = \begin{cases} 
A e^{K_1z} & z < 0 \\
Be^{K_3z} + Ce^{-K_3z} & 0 \leq z \leq d \\
D e^{-K_2z} & z < d
\end{cases}, \quad (2.241)
\]

where [see Eq. (2.238)]
\[
K_l = \sqrt{\kappa^2 - \left( \frac{m\omega}{c} \right)^2}. \quad (2.242)
\]

In a matrix form the boundary conditions at the interfaces \(z = 0\) and \(z = d\) can be expressed as
\[
M_s \begin{pmatrix} A \\ B \\ C \\ D \end{pmatrix}^T = 0, \quad (2.243)
\]

where
\[
M_s = \begin{pmatrix}
1 & -1 & 0 \\
0 & e^{K_3d} & -e^{-K_3d} \\
\frac{K_1}{\mu_1} & -\frac{K_3}{\mu_3} & \frac{K_3}{\mu_3} \\
0 & \frac{K_2 e^{K_3d}}{\mu_3} & -\frac{K_2 e^{-K_3d}}{\mu_2}
\end{pmatrix}. \quad (2.244)
\]
Chapter 2. Maxwell’s Equations in Matter

A nontrivial solutions exists provided that

\[ 0 = \det M_s = e^{-2K_3d} \left( \frac{\bar{\mu}_s}{K_3} - \frac{\mu_1}{K_1} \right) \left( \frac{\mu_3}{K_3} - \frac{\bar{\mu}_2}{K_2} \right)^2 A_s, \]

(2.245)

where

\[ A_s = e^{2K_3d} \left( \frac{\bar{\mu}_s}{K_3} \right) \left( \frac{\mu_1}{K_1} \right) \left( \frac{\mu_3}{K_3} \right) - 1. \]

(2.246)

Thus the angular frequencies \( \omega \) associated with s-polarization surface modes can be found by solving the equation \( A_s = 0 \). Similarly, for the case of p-polarization, the solution for the magnetic field \( H \) for each layer is assumed to have the form

\[ H = \hat{y} h_y (z) e^{i\kappa \cdot \rho}, \]

(2.247)

where [see Eq. (2.153)]

\[ \frac{h_y'}{h_y} = K^2. \]

(2.248)

The electric field \( E \) is given by [see Eq. (2.140)]

\[ E = i \frac{k_e}{\kappa_0 \epsilon} \nabla \times H = - \frac{(-h_y', 0, 0) + i (0, 0, k_x h_y)}{ik_0 \epsilon} e^{i\kappa \cdot \rho}. \]

(2.249)

Thus, for p-polarization the boundary condition (2.79) implies that \( h_y \) is continuous, and the boundary condition (2.80) implies that \( \epsilon^{-1} h_y' \) is continuous. In a similar fashion one finds that the angular frequencies \( \omega \) associated with p-polarization surface modes can be calculated by solving the equation \( A_p = 0 \), where

\[ A_p = e^{2K_3d} \left( \frac{\bar{\mu}_p}{K_3} \right) \left( \frac{\mu_1}{K_1} \right) \left( \frac{\mu_3}{K_3} \right) - 1. \]

(2.250)

6. For the above discussed trilayer, the Casimir force between layers 1 and 2 is calculated below by summing over all angular frequencies \( \omega \) corresponding to surface modes with either s-polarization or p-polarization [see Eqs. (2.242), (2.246) and (2.250)].
a) The Casimir pressure (force per unit area) $P(d)$ is given by

$$P(d) = -\frac{\partial u(d)}{\partial d}, \quad (2.251)$$

where $u(d)$ is the zero point energy per unit area associated with the surface modes, which is found by summing over all angular frequencies corresponding to both s-polarization and p-polarization, and multiplying the sum by $\hbar/2A$, where $A$ is the area. The summation over allowed values of $k_x$ and $k_y$ can be performed using the rule

$$\sum_{k_x,k_y} \sim -\frac{A}{4\pi^2} \int_{-\infty}^{\infty} dk_x \int_{-\infty}^{\infty} dk_y \rightarrow \frac{A}{2\pi} \int_0^\infty d\kappa \kappa. \quad (2.252)$$

For a given $\kappa$ the angular frequencies $\omega$ can be found by solving $A_s(\omega,\kappa) = 0$ and $A_p(\omega,\kappa) = 0$. This can be performed with the help of the argument theorem

$$P(d) = -\frac{\hbar}{4\pi} \frac{\partial}{\partial d} \int_0^\infty d\kappa \kappa T(\kappa), \quad (2.253)$$

where

$$T(\kappa) = \frac{1}{2\pi i} \oint_C \omega \left(\frac{\partial \log A_s}{\partial \omega} + \frac{\partial \log A_p}{\partial \omega}\right), \quad (2.254)$$

and where the integration contour $C$ in the $\omega$ complex plane is assumed to enclose all zeros of $A_s$ and $A_p$. The contour $C$ is chosen to contain a section along the imaginary axis from $\omega = -iR$ to $\omega = iR$ and a semi-circle of radius $R$ in the real positive half complex plane (i.e. right to the imaginary axis). In the limit $R \to \infty$ the integral along the semi-circle vanishes. By integrating along the imaginary axis and by performing integration by parts Eq. (2.253) becomes

$$P(d) = \frac{\hbar}{8\pi^2} \frac{\partial}{\partial d} \int_0^\infty d\kappa \kappa \int_{-\infty}^{\infty} d\Omega \left(\log A_s + \log A_p\right), \quad (2.255)$$

where $\omega = i\Omega$, thus (note that the integrand is an even function of $\Omega$)

$$P(d) = -\frac{\hbar}{2\pi^2} \int_0^\infty d\kappa \kappa \int_0^\infty d\Omega \left(\frac{K_3(A_s+1)}{A_s} + \frac{K_3(A_p+1)}{A_p}\right)$$

$$= -\frac{\hbar}{\pi^2} \int_0^\infty d\kappa \kappa \int_0^\infty d\Omega K_3$$

$$- \frac{\hbar}{2\pi^2} \int_0^\infty d\kappa \kappa \int_0^\infty d\Omega K_3 \left(\frac{1}{A_s} + \frac{1}{A_p}\right). \quad (2.256)$$
The first term is independent on \( \epsilon_1 \) and \( \epsilon_2 \). After disregarding it Eq. (2.256) becomes

\[
P(d) = -\frac{\hbar}{2\pi^2} \int_0^\infty \mathrm{d} \kappa \kappa \int_0^\infty \mathrm{d} \Omega K_3 \left( \frac{1}{A_s} + \frac{1}{A_p} \right).
\]  

(2.257)

The variable \( p \) is defined by

\[
\kappa^2 = \frac{n_3^2 \Omega^2}{c^2} (p^2 - 1),
\]

(2.258)

the variables \( s_{1,2} \) by

\[
s_{1,2} = \sqrt{p^2 - 1 + \frac{n_{1,2}^2}{n_3^2}},
\]

(2.259)

and the variable \( x \) by

\[
x = 2K_3d.
\]

(2.260)

The following holds (recall that \( \omega = i\Omega \))

\[
K_1 = \frac{n_3 \Omega}{c} \sqrt{p^2 - 1 + \frac{n_1^2}{n_3^2}},
\]

(2.261)

\[
K_2 = \frac{n_3 \Omega}{c} \sqrt{p^2 - 1 + \frac{n_2^2}{n_3^2}},
\]

(2.262)

\[
K_3 = \frac{n_3 \Omega}{c} p,
\]

(2.263)

or

\[
K_1 = \frac{x}{2pd} \sqrt{p^2 - 1 + \frac{n_1^2}{n_3^2}},
\]

(2.264)

\[
K_2 = \frac{x}{2pd} \sqrt{p^2 - 1 + \frac{n_2^2}{n_3^2}},
\]

(2.265)

\[
K_3 = \frac{x}{2d}.
\]

(2.266)

By employing the above definitions and relations one finds that Eq. (2.256) can be expressed as [see Eqs. (2.246) and (2.250)]

\[
P(d) = -\frac{\hbar c}{32\pi^2 d^4} \int_1^\infty \frac{\mathrm{d} p}{p^2} \int_0^\infty \mathrm{d} x x^3 n_3^{-1} \\
\times \left( \frac{1}{\zeta_{s,1}s_{s,2}e^x - 1} + \frac{1}{\zeta_{p,1}\zeta_{p,2}e^x - 1} \right),
\]

(2.267)
2.13. Solutions

where

$$\zeta_{s,n} = \frac{\mu_n K_n + \mu_3 K_n}{\mu_n K_n - \mu_3 K_n},$$  \hspace{1cm} (2.268)

$$\zeta_{p,n} = \frac{\epsilon_n K_n + \epsilon_3 K_n}{\epsilon_n K_n - \epsilon_3 K_n},$$ \hspace{1cm} (2.269)

and $n \in \{1, 2\}$.

b) For this case [recall that $\omega = i\Omega$ and see Eqs. (2.264), (2.265) and (2.266)]

$$\zeta_{s,n} = 1 + \sqrt{1 + x^{-2} \left( \frac{d}{d\omega} \right)^2 \frac{1}{1 - \sqrt{1 + x^{-2} \left( \frac{d}{d\omega} \right)^2}}},$$ \hspace{1cm} (2.270)

$$\zeta_{p,n} = 1 + x^{-2} \left( \frac{d}{d\omega} \right)^2 + \sqrt{1 + x^{-2} \left( \frac{d}{d\omega} \right)^2 \frac{1}{1 + x^{-2} \left( \frac{d}{d\omega} \right)^2}},$$ \hspace{1cm} (2.271)

where

$$d_p = \frac{c}{2\omega_p}.$$ \hspace{1cm} (2.272)

For the present case the following holds [see Eqs. (2.270) and (2.271)]

$$1 - \frac{1}{\zeta_{s,n} e^x - 1} \frac{\zeta_{p,n} e^x - 1}{e^x - 1} = \frac{2}{\zeta_{p,n} e^x - 1}$$

$$- \frac{4xe^x}{(e^x - 1)^2} \left( 1 + \frac{1}{p^2} \right) \frac{d_p}{d} + O \left( \left( \frac{d_p}{d} \right)^2 \right),$$ \hspace{1cm} (2.273)

thus when $d_p \ll d$ the Casimir force (2.267) is approximately given by

$$P (d) = P_{pc} (d) + P_{fcc} (d),$$ \hspace{1cm} (2.274)

where the term $P_{pc} (d)$, which is given by

$$P_{pc} (d) = \frac{\hbar c}{16\pi^2 d^4} \int_1^\infty dp \int_0^{\infty} x^3 dx \frac{x^3}{e^x - 1},$$

$$= -\frac{\pi^4 \hbar c}{240\pi^2 d^4},$$ \hspace{1cm} (2.275)
represents the force in the limit of infinite conductivity, and where the term \( P_{\text{fcc}}(d) \), which is given by

\[
P_{\text{fcc}}(d) = -\frac{32}{3} d^3 P_{\text{pc}}(d),
\]  

(2.276)

is the correction due to finite conductivity. The above results are obtained using the following identities

\[
\int_0^\infty \frac{x^3}{e^x - 1} \, dx = \frac{\pi^4}{15},
\]

(2.277)

\[
\int_1^\infty \frac{dp}{p^2} = 1,
\]

(2.278)

\[
\int_0^\infty \frac{x^4 e^x}{(e^x - 1)^2} \, dx = \frac{4\pi^4}{15},
\]

(2.279)

\[
\int_1^\infty \frac{1 + \frac{1}{p^2}}{p^2} \, dp = \frac{4}{3}.
\]

(2.280)

7. The transformed electric field \( \mathbf{E}' = (E'_x, E'_y, E'_z) \) is given by [see Eq. (2.35)]

\[
\mathbf{E}' = \mathbf{E} + \gamma (E_{\perp} + \hat{\beta} \times \mathbf{B}_{\perp}) = \gamma \mathbf{E} + (1 - \gamma) \left( \mathbf{E} \cdot \hat{\beta} \right) \hat{\beta} + \gamma \beta \hat{\beta} \times \mathbf{B},
\]

(2.281)

where \( \gamma = 1/\sqrt{1 - \beta^2} \) and \( \mathbf{B} = \hat{z} \times \mathbf{E} = (-E_y, E_x, 0) \) [see Eq. (2.141)], and thus for the case where \( \hat{\beta} = (\sin \theta, 0, \cos \theta) \) one has

\[
E'_x = (1 + (\gamma - 1) \cos^2 \theta - \beta \gamma \cos \theta) E_x,
\]

(2.282)

\[
E'_y = \gamma (1 - \beta \cos \theta) E_y,
\]

(2.283)

\[
E'_z = ((1 - \gamma) \cos \theta + \beta \gamma) E_x \sin \theta.
\]

(2.284)

Let \( \hat{\mathbf{n}}' \) be a unit vector in the direction of propagation of the wave as measured by the moving observer. With the help of the aberration of light formula (1.61) one finds that

\[
\hat{\mathbf{n}}' = \hat{\mathbf{n}} - \gamma \left( \frac{1 - \frac{\gamma}{1 + \gamma} (\hat{\beta} \cdot \hat{\mathbf{n}})}{\gamma (1 - (\hat{\beta} \cdot \hat{\mathbf{n}}))} \right) \hat{\beta}
\]

(2.285)

where for the current case \( \hat{\mathbf{n}} = \hat{z} \). The vector \( \hat{\mathbf{n}}' \) lie in the \( xz \) plane and it makes an angle \( \alpha \) with the \( z \) axis given by [note that \( \beta^2 = (1 + \gamma^{-1}) (1 - \gamma^{-1}) \)]

\[
\tan \alpha = -\frac{((1 - \gamma) \cos \theta + \beta \gamma) \sin \theta}{1 + (\gamma - 1) \cos^2 \theta - \beta \gamma \cos \theta},
\]

(2.286)
2.13. Solutions

and the following holds \( \tan \alpha = -E'_z/E'_x \) [see Eqs. (2.282) and (2.284)]. Rotation of \( \mathbf{E}' \) about the \( y \) axis by an angle \( \alpha \) leads to (note that the rotation preserves both the \( y \) component and the length)

\[
\begin{pmatrix}
E'_x \\
E'_y \\
E'_z
\end{pmatrix}
\rightarrow
\begin{pmatrix}
E'_{xR} \\
E'_{yR} \\
E'_{zR}
\end{pmatrix}
= \begin{pmatrix}
\sqrt{E'^2_x + E'^2_z} \\
0
\end{pmatrix},
\]

(2.287)

and thus

\[
\begin{pmatrix}
E'_{xR} \\
E'_{yR}
\end{pmatrix}
= \gamma (1 - \beta \cos \theta) \begin{pmatrix}
E_x \\
E_y
\end{pmatrix},
\]

(2.288)

and therefore the Stokes parameters are transformed according to

\[
\begin{pmatrix}
S'_0 \\
S'_1 \\
S'_2 \\
S'_3
\end{pmatrix}
= \gamma^2 (1 - \beta \cos \theta)^2
\begin{pmatrix}
S_0 \\
S_1 \\
S_2 \\
S_3
\end{pmatrix}.
\]

(2.289)

As can be seen from Eq. (1.168), the factor \( \gamma (1 - \beta \cos \theta) \) is the frequency ratio \( \omega'/\omega \) due to the Doppler effect.

8. The following holds [see Eqs. (1.21), (2.32) and (2.49)]

\[
\eta \hat{F} \eta \hat{F} = \Lambda^{-1} \eta \hat{F}' \eta \hat{F}' \Lambda,
\]

(2.290)

\[
\eta \hat{G} \eta \hat{G} = \Lambda^{-1} \eta \hat{G}' \eta \hat{G}' \Lambda,
\]

(2.291)

and [see Eqs. (1.14), (2.29) and (2.46)]

\[
\frac{1}{2} \text{Tr} \left( \eta \hat{F} \eta \hat{F} \right) = E^2 - B^2,
\]

(2.292)

\[
\frac{1}{2} \text{Tr} \left( \eta \hat{G} \eta \hat{G} \right) = D^2 - H^2.
\]

(2.293)

In general, for any two square matrices \( M_1 \) and \( M_2 \) the following holds

\[
\text{Tr} (M_1 M_2) = \text{Tr} (M_2 M_1),
\]

and thus

\[
\text{Tr} \left( \eta \hat{F} \eta \hat{F} \right) = \text{Tr} \left( \eta \hat{F}' \eta \hat{F}' \right),
\]

(2.294)

\[
\text{Tr} \left( \eta \hat{G} \eta \hat{G} \right) = \text{Tr} \left( \eta \hat{G}' \eta \hat{G}' \right).
\]

(2.295)

i.e. both \( E^2 - B^2 \) and \( D^2 - H^2 \) are Lorentz invariant. The following holds [see Eqs. (2.35), (2.36) and (3.265)]

\[
\mathbf{E} \cdot \mathbf{B} = E'_\parallel \cdot B'_\parallel + \gamma^2 (E'_\perp - \beta \times B'_\perp) \cdot (B'_\perp + \beta \times E'_\perp)
\]

\[
= E'_\parallel \cdot B'_\parallel + \gamma^2 (1 - \beta^2) E'_\perp \cdot B'_\perp + \gamma^2 (E'_\perp \cdot (\beta \times E'_\perp) - (\beta \times B'_\perp) \cdot B'_\perp)
\]

\[
= E' \cdot B',
\]

(2.296)
and thus $\mathbf{E} \cdot \mathbf{B}$ is Lorentz invariant. In a similar way one finds that $\mathbf{D} \cdot \mathbf{H}$ is Lorentz invariant [see Eqs. (2.106) and (2.107)]. Moreover, with the help of the general identity

\[ (\mathbf{A} \times \mathbf{B}) \cdot \mathbf{C} = (\mathbf{B} \times \mathbf{C}) \cdot \mathbf{A} = (\mathbf{C} \times \mathbf{A}) \cdot \mathbf{B}, \]

(2.297)

one finds that [see Eqs. (2.35), (2.36), (2.106) and (2.107)]

\[ \mathbf{E} \cdot \mathbf{D} - \mathbf{B} \cdot \mathbf{H} = \mathbf{E}' \cdot \mathbf{D}' - \mathbf{B}' \cdot \mathbf{H}' + \gamma^2 (\mathbf{E}' \perp - \beta \times \mathbf{B}' \perp) \cdot (\mathbf{D}' \perp - \beta \times \mathbf{H}' \perp) \]

\[ - \gamma^2 (\mathbf{B}' \perp + \beta \times \mathbf{E}' \perp) \cdot (\mathbf{H}' \perp + \beta \times \mathbf{D}' \perp) \]

\[ = \mathbf{E}' \cdot \mathbf{D}' - \mathbf{B}' \cdot \mathbf{H}', \]

(2.298)

and thus $\mathbf{E} \cdot \mathbf{D} - \mathbf{B} \cdot \mathbf{H}$ is Lorentz invariant.

9. In the Lorenz gauge the Maxwell’s equations in vacuum can be expressed as [see Eqs. (2.62) and (2.63)]

\[ \Box^2 \mathbf{A} = - \frac{4\pi}{c^2} \mathbf{J}, \]

(2.299)

where $\Box^2 = -c^{-2} \partial^2 / \partial t^2 + \nabla^2$ is the D’Alembertian operator, $\mathbf{A} = (\phi, A_1, A_2, A_3)^T$ is the potential 4-vector and $\mathbf{J} = (c\rho, J_1, J_2, J_3)^T$ is the current 4-vector. Both $\mathbf{A}(t, \mathbf{x})$ and $\mathbf{J}(t, \mathbf{x})$ can be Fourier expanded (in time only) as

\[ \mathbf{A}(t, \mathbf{x}) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} d\omega \, A(\omega, \mathbf{x}) e^{i\omega t}, \]

(2.300)

\[ \mathbf{J}(t, \mathbf{x}) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} d\omega \, J(\omega, \mathbf{x}) e^{i\omega t}. \]

(2.301)

Substituting into Eq. (2.299) yields

\[ \left( \frac{\omega^2}{c^2} + \nabla^2 \right) \mathbf{A}(\omega, \mathbf{x}) = - \frac{4\pi}{c} \mathbf{J}(\omega, \mathbf{x}). \]

(2.302)

As is shown below in chapter 5 [see Eqs. (5.84) and (5.93)], the following holds

\[ (k^2 + \nabla^2) \mathbf{g}(\mathbf{x} - \mathbf{x}') = \delta(\mathbf{x} - \mathbf{x}'), \]

(2.303)

where $k$ is a constant, and the so-called Green function $\mathbf{g}(\mathbf{x} - \mathbf{x}')$ is given by

\[ \mathbf{g}(\mathbf{x} - \mathbf{x}') = -\frac{e^{\pm ik|\mathbf{x} - \mathbf{x}'|}}{4\pi |\mathbf{x} - \mathbf{x}'|}. \]

(2.304)
Thus the solution of Eq. (2.302) can be expressed as

\[ A(\omega, x) = \int d^3x' \frac{e^{\pm i\omega |x-x'|}}{c|x-x'|} J(\omega, x'). \]  
(2.305)

Applying the inverse Fourier transform in time leads to [see Eq. (5.6)]

\[ A(t, x) = \int d^3x' \int_0^\infty dt' \frac{J(t', x')}{c|x-x'|} 2\pi \int_0^\infty d\omega e^{i\omega(t \pm |x-x'|/c - t')} \delta(t \pm |x-x'|/c - t') \]

\[ = \int d^3x' \frac{J(t \pm |x-x'|/c, x')}{c|x-x'|}, \]  
(2.306)

Due to the principle of causality, the solution with the + sign is rejected. The potential with the minus sign, which is given by Eq. (2.187), is commonly called the retarded potential. As is expected from the principle of causality, propagation at the speed of light results in the value of \( J \) at time \( t - |x-x'|/c \) and location \( x' \) affecting the value of \( A \) at time \( t \) and location \( x \).
3. Geometrical Optics

In the theory of geometrical optics the Maxwell’s equations are simplified based on the assumption that the characteristic wavelength $\lambda$ of electromagnetic waves can be considered as small (in comparison with relevant length scales in the problem under study).

3.1 Scalar Geometrical Optics

In this section the short wavelength approximation is demonstrated for the relatively simple case of a scalar field. Consider the following scalar wave equation [compare with Eq. (3.192)]

$$\left( \frac{n^2(r)}{c^2} \frac{\partial^2}{\partial t^2} - \nabla^2 \right) U(r, t) = 0 . \quad (3.1)$$

By substituting a solution having the form

$$U(r, t) = u(r) e^{-i\omega t} , \quad (3.2)$$

into Eq. (3.1) one finds that $u(r)$ satisfies the following Helmholtz equation [see Eq. (2.153)]

$$(\nabla^2 + n^2 k_0^2) u(r) = 0 , \quad (3.3)$$

where

$$k_0 = \frac{\omega}{c} . \quad (3.4)$$

Consider a solution for $u(r)$ having the form

$$u(r) = u_E(r) e^{ik_0 \psi(r)} , \quad (3.5)$$

where $u_E(r)$ is expressed as an asymptotic expansion in powers of $1/k_0$

$$u_E(r) = u_0(r) + \frac{1}{k_0} u_1(r) + \frac{1}{k_0^2} u_2(r) + \cdots , \quad (3.6)$$

and where the real function $\psi(r)$ is called the eikonal (image in greek). In geometrical optics the parameter $1/k_0$ is assumed small.
Chapter 3. Geometrical Optics

Claim. To first order in $1/k_0$ the following holds

$$(\nabla \psi)^2 = n^2 ,$$

and

$$u_0 \nabla^2 \psi + 2 (\nabla \psi) \cdot (\nabla u_0) = 0 .$$

Proof. By substituting Eq. (3.5) into Eq. (3.3) and employing the general identity

$$\nabla^2 (fg) = f \nabla^2 g + g \nabla^2 f + 2 (\nabla f) \cdot (\nabla g) ,$$

one obtains

$$k_0^{-2} \nabla^2 u_E + ik_0^{-1} u_E (\nabla^2 \psi + ik_0 (\nabla \psi)^2) + 2ik_0^{-1} (\nabla \psi) \cdot (\nabla u_E) + n^2 u_E = 0 .$$

Collecting all terms of zeroth order in $1/k_0$ yields

$$[n^2 - (\nabla \psi)^2] u_0 = 0 ,$$

and collecting all terms of 1'st order in $1/k_0$ yields

$$i [u_0 \nabla^2 \psi + 2 (\nabla \psi) \cdot (\nabla u_0)] + [n^2 - (\nabla \psi)^2] u_1 = 0 .$$

Unless $u_0 (r)$ vanishes everywhere Eq. (3.11) leads to Eq. (3.7), which is called the eikonal equation. By employing the eikonal equation one finds that Eq. (3.12) becomes Eq. (3.8). Note that multiplying Eq. (3.8) by $u_0$ leads to

$$u_0^2 \nabla^2 \psi + (\nabla \psi) \cdot (\nabla u_0^2) = 0 ,$$

thus Eq. (3.8) can be rewritten as

$$\nabla \cdot (u_0^2 \nabla \psi) = 0 .$$

3.2 Vectorial Geometrical Optics

As has been shown in the previous section, the eikonal equation (3.7) can be derived from the scalar approximation. However, the vectorial nature of electromagnetism requires a vectorial analysis. The starting point of such analysis is the version of the Maxwell’s equations given by Eqs. (2.140), (2.141), (2.142) and (2.143)
3.2. Vectorial Geometrical Optics

\[ \nabla \times \mathbf{H} = -i k_0 \epsilon \mathbf{E}, \quad (3.15) \]
\[ \nabla \times \mathbf{E} = i k_0 \mu \mathbf{H}, \quad (3.16) \]
\[ \nabla \cdot (\epsilon \mathbf{E}) = 0, \quad (3.17) \]
\[ \nabla \cdot (\mu \mathbf{H}) = 0, \quad (3.18) \]

where
\[ k_0 = \frac{\omega}{c}. \quad (3.19) \]

Recall that Eqs. (2.140), (2.141), (2.142) and (2.143) have been derived based on the assumptions that the medium is inhomogeneous, free of sources, isotropic, linear and stationary, the conductivity \( \sigma \) vanishes, and \( \epsilon = \epsilon (\mathbf{r}) \) and \( \mu = \mu (\mathbf{r}) \) are taken to be time independent scalars. In addition, harmonic time dependency has been assumed.

### 3.2.1 Asymptotic Expansion

In the so-called Luneberg-Kline asymptotic expansion \( \mathbf{E} \) and \( \mathbf{H} \) are expressed in terms of the eikonal function \( \psi (\mathbf{r}) \) as

\[ \mathbf{E} (\mathbf{r}) = \exp \{ i k_0 \psi (\mathbf{r}) \} \sum_{m=0}^{\infty} \frac{\mathbf{E}_m (\mathbf{r})}{(i k_0)^m}, \quad (3.20) \]

and

\[ \mathbf{H} (\mathbf{r}) = \exp \{ i k_0 \psi (\mathbf{r}) \} \sum_{m=0}^{\infty} \frac{\mathbf{H}_m (\mathbf{r})}{(i k_0)^m}. \quad (3.21) \]

**Exercise 3.2.1.** Show that

\[ \nabla \times \mathbf{H}_m = -\epsilon \mathbf{E}_{m+1} - (\nabla \psi) \times \mathbf{H}_{m+1}, \quad (3.22) \]
\[ \nabla \times \mathbf{E}_m = \mu \mathbf{H}_{m+1} - (\nabla \psi) \times \mathbf{E}_{m+1}, \quad (3.23) \]
\[ \nabla \cdot (\epsilon \mathbf{E}_m) = -\epsilon \mathbf{E}_{m+1} \cdot \nabla \psi, \quad (3.24) \]
\[ \nabla \cdot (\mu \mathbf{H}_m) = -\mu \mathbf{H}_{m+1} \cdot \nabla \psi. \quad (3.25) \]

**Solution 3.2.1.** Substituting Eqs. (3.20) and (3.21) into Eqs. (2.140), (2.141), (2.142) and (2.143) leads to

\[ \sum_{m=0}^{\infty} \frac{1}{(i k_0)^m} \nabla \times [\mathbf{H}_m \exp \{ i k_0 \psi \}] = -\epsilon \exp \{ i k_0 \psi \} \sum_{m=0}^{\infty} \frac{1}{(i k_0)^m} \mathbf{E}_{m+1}, \quad (3.26) \]
\[ \sum_{m=0}^{\infty} \frac{1}{(i k_0)^m} \nabla \times [\mathbf{E}_m \exp \{ i k_0 \psi \}] = \mu \exp \{ i k_0 \psi \} \sum_{m=0}^{\infty} \frac{1}{(i k_0)^m} \mathbf{H}_{m+1}, \quad (3.27) \]
\[
\sum_{m=0}^{\infty} \frac{1}{(ik_0)^m} \nabla \cdot [\epsilon E_m \exp (ik_0 \psi)] = 0,
\]
(3.28)

and

\[
\sum_{m=0}^{\infty} \frac{1}{(ik_0)^m} \nabla \cdot [\mu H_m \exp (ik_0 \psi)] = 0.
\]
(3.29)

By using the vectors identities (2.151) and (2.152) one obtains

\[
\sum_{m=0}^{\infty} \frac{1}{(ik_0)^m} \nabla \times H_m = \sum_{m=-1}^{\infty} \frac{1}{(ik_0)^m} [\epsilon E_{m+1} \cdot (\nabla \psi) \times H_{m+1}],
\]
(3.30)

\[
\sum_{m=0}^{\infty} \frac{1}{(ik_0)^m} \nabla \times E_m = \sum_{m=-1}^{\infty} \frac{1}{(ik_0)^m} [\mu H_{m+1} \cdot (\nabla \psi) \times E_{m+1}],
\]
(3.31)

\[
\sum_{m=0}^{\infty} \frac{1}{(ik_0)^m} \nabla \cdot (\epsilon E_m) + \sum_{m=-1}^{\infty} \frac{1}{(ik_0)^m} E_{m+1} \cdot \nabla \psi = 0,
\]
(3.32)

\[
\sum_{m=0}^{\infty} \frac{1}{(ik_0)^m} \nabla \cdot (\mu H_m) + \sum_{m=-1}^{\infty} \frac{1}{(ik_0)^m} H_{m+1} \cdot \nabla \psi = 0,
\]
(3.33)

in agreement with Eqs. (3.22), (3.23), (3.24) and (3.25).

### 3.2.2 Eikonal Equation

By substituting the Luneberg-Kline asymptotic expansion (3.20) and (3.21) into the Maxwell’s equations (2.140), (2.141), (2.142) and (2.143) while keeping only \(k_0\) independent terms one obtains

\[
\epsilon E_0 + (\nabla \psi) \times H_0 = 0,
\]
(3.34)

\[
\mu H_0 - (\nabla \psi) \times E_0 = 0,
\]
(3.35)

\[
E_0 \cdot \nabla \psi = 0,
\]
(3.36)

\[
H_0 \cdot \nabla \psi = 0.
\]
(3.37)

Equations (3.34) and (3.35) imply that

\[
n^2 E_0 + (\nabla \psi) \times [(\nabla \psi) \times E_0] = 0,
\]
(3.38)

thus, by using the vector identity (1.95), which is given by

\[
A \times (B \times C) = (A \cdot C) B - (A \cdot B) C,
\]
(3.39)

together with Eq. (3.36) one obtains

\[
\left[ n^2 - (\nabla \psi)^2 \right] E_0 = 0.
\]
(3.40)

The assumption that \(E_0\) does not vanish everywhere leads to the eikonal equation [see Eq. (3.7)]

\[
(\nabla \psi)^2 = n^2.
\]
(3.41)
### 3.3 Optical Rays

The optical rays are defined as trajectories orthogonal to the wave front surfaces of constant $\psi$ (wave front surfaces). Let $\mathbf{r}(s)$ be an optical ray with arc-length parametrization, namely $d\mathbf{r}/ds = \hat{s}$, where $\hat{s}$ is a unit vector. The ray equation reads

$$\frac{d\mathbf{r}}{ds} = \frac{\nabla \psi}{n} = \hat{s}. \quad (3.42)$$

The normal unit vector $\hat{\nu}$ and the curvature $\kappa$ are defined by the relation

$$\frac{d\hat{s}}{ds} = \kappa \hat{\nu}. \quad (3.43)$$

One can easily show that $\hat{\nu} \cdot \hat{s} = 0$ by taking the derivative of the relation $\hat{s} \cdot \hat{s} = 1$ with respect to $s$. The vectors $\hat{s}$, $\hat{\nu}$ together with the binormal unit vector $\hat{b}$, which is defined by $\hat{b} = \hat{s} \times \hat{\nu}$, form a local orthonormal coordinate frame.

*Claim.* The following holds

$$\frac{d}{ds} \begin{pmatrix} \hat{s} \\ \hat{\nu} \\ \hat{b} \end{pmatrix} = \begin{pmatrix} 0 & \kappa & 0 \\ -\kappa & 0 & \tau \\ 0 & \tau & 0 \end{pmatrix} \begin{pmatrix} \hat{s} \\ \hat{\nu} \\ \hat{b} \end{pmatrix}, \quad (3.44)$$

where $\tau$ is real.

*Proof.* By taking the derivative of $\hat{s} \cdot \hat{\nu} = 0$ with respect to $s$ one finds that

$$\hat{s} \cdot \frac{d\hat{\nu}}{ds} = -\kappa. \quad (3.45)$$

Similarly, by taking the derivative of $\hat{b} \cdot \hat{\nu} = 0$ with respect to $s$ one obtains

$$\hat{b} \cdot \frac{d\hat{\nu}}{ds} = -\hat{\nu} \cdot \frac{d\hat{b}}{ds}. \quad (3.46)$$

By employing the definition $\hat{b} = \hat{s} \times \hat{\nu}$ and Eq. (3.43) one finds that

$$\frac{d\hat{b}}{ds} = \hat{s} \times \frac{d\hat{\nu}}{ds}, \quad (3.47)$$

thus

$$\hat{s} \cdot \frac{d\hat{b}}{ds} = 0. \quad (3.48)$$

Moreover, by taking the derivative of $\hat{b} \cdot \hat{b} = 1$ with respect to $s$ one finds that
\[ \hat{b} \cdot \frac{d\hat{b}}{ds} = 0 , \quad (3.49) \]

thus \( \frac{d\hat{b}}{ds} \) is parallel to \( \hat{\nu} \). The torsion \( \tau \) is defined as
\[ \frac{d\hat{b}}{ds} = -\tau \hat{\nu} . \quad (3.50) \]

The above results (and definitions) can be summarized by Eq. (3.44).

Alternatively, Eq. (3.44) can be expressed as
\[ \frac{d}{ds} \begin{pmatrix} \hat{s} \\ \hat{\nu} \\ \hat{b} \end{pmatrix} = \delta \times \begin{pmatrix} \hat{s} \\ \hat{\nu} \\ \hat{b} \end{pmatrix} , \quad (3.51) \]

where
\[ \delta = \tau \hat{s} + \kappa \hat{b} . \quad (3.52) \]

**Exercise 3.3.1.** Show that for any closed curve \( C \) the following holds
\[ \oint_C (n\hat{s}) \cdot dl = 0 . \quad (3.53) \]

**Solution 3.3.1.** With the help of Eq. (3.42) one finds that
\[ \nabla \times (n\hat{s}) = \nabla \times \nabla \psi = 0 , \quad (3.54) \]

and thus according to Stoke’s theorem [see Eq. (2.67)] Eq. (3.53), which is called the Lagrange’s integral invariant, holds.

### 3.3.1 Reflection and Refraction

Consider an optical ray striking the interface between two homogeneous materials of refraction indices \( n_1 \) and \( n_2 \). Part of the ray is reflected and part is refracted. The angles between the incident, reflected and refracted rays and the normal to the interface are denoted as \( \theta_i \), \( \theta_r \) and \( \theta_t \), respectively (see Fig. 3.1).

By assuming that the theory of geometrical optics is applicable for the case of abrupt change in \( n \) at the interface between two different materials, one can obtain a relation between the angles \( \theta_i \), \( \theta_r \) and \( \theta_t \). By employing the Lagrange’s integral invariant (3.53), which implies that the tangential component of \( n\hat{s} \) is continuous [compare with Eq. (2.80)], one obtains the relation
\[ \theta_i = \theta_r , \]

and the so-called Snell’s law, which is given by [compare with Eq. (2.224)]
\[ n_1 \sin \theta_i = n_2 \sin \theta_t . \quad (3.55) \]
3.3. Optical Rays

![Diagram of optical rays](image)

**Fig. 3.1.** Incident, reflected and refracted rays.

### 3.3.2 The Ray Equation

A ray can be traced by solving the ray equation, which is given by Eq. (3.56) below.

**Claim.** The following holds

\[
\kappa \hat{r} = \nabla (\log n) - \hat{s} \frac{d}{ds} (\log n)
\]

(3.56)

**Proof.** Taking the derivative \(d/ds\) of Eq. (3.42) leads to

\[
\frac{d}{ds} \left( \frac{n}{ds} \right) = \frac{d}{ds} \nabla \psi = \frac{1}{n} \left( \nabla \psi \cdot \nabla \right) \nabla \psi .
\]

(3.57)

By using the vector identities

\[
\nabla (A^2) = 2 \left[ A \times (\nabla \times A) + (A \cdot \nabla) A \right],
\]

(3.58)

and

\[
\nabla \times (\nabla \psi) = 0 ,
\]

(3.59)

one finds that

\[
\frac{d}{ds} \left( \frac{n}{ds} \right) = \frac{1}{2n} \nabla (\nabla \psi)^2 ,
\]

(3.60)

or [see Eq. (3.41)]

\[
\frac{d}{ds} \left( \frac{n}{ds} \right) = \frac{1}{2n} \nabla n^2 = \nabla n .
\]

(3.61)
Moreover, the following holds
\[
\frac{d}{ds} \left( \frac{n}{ds} \right) = \hat{s} \frac{dn}{ds} + n \kappa \hat{\nu} .
\] (3.62)

The last two results directly lead to Eq. (3.56).

By multiplying Eq. (3.56) by \( \hat{\nu} \) one obtains
\[
\kappa = \hat{\nu} \cdot \nabla (\log n) ,
\] (3.63)
thus the ray is curved towards the regime of higher \( n \). One also finds that
the vector \( \nabla n \) is in the so-called osculating plane of the ray \( (\hat{s} \hat{\nu} \) plane).

**Exercise 3.3.2.** Show that
\[
\frac{d\hat{s}}{ds} = \hat{s} \times (\nabla (\log n) \times \hat{s}) .
\] (3.64)

**Solution 3.3.2.** The general identity [see Eq. (1.95)]
\[
A \times (B \times A) = (A \cdot A) B - (A \cdot B) A ,
\] (3.65)
together with Eq. (3.56) lead to Eq. (3.64).

**Exercise 3.3.3.** Consider a parametrization of an optical ray given by
\[
r (s) = r (\sigma (s)) ,
\] (3.66)
where
\[
\frac{ds}{d\sigma} = n .
\] (3.67)

Show that
\[
\frac{d^2 r}{d\sigma^2} = -\nabla U ,
\] (3.68)
where
\[
U = \frac{n^2}{2} .
\] (3.69)

**Solution 3.3.3.** With the help of Eqs. (3.61) and the relation [see (3.67)]
\[
\frac{d}{ds} = \frac{1}{n} \frac{d}{d\sigma} ,
\] (3.70)
on one finds that
\[
\frac{d^2 r}{d\sigma^2} = \nabla \left( \frac{n^2}{2} \right) .
\] (3.71)
The above result (3.68) demonstrate the analogy between geometrical optics and mechanics (Newton’s second law).

**Exercise 3.3.4.** Consider a spherically symmetric medium, for which

\[ n = n(r), \tag{3.72} \]

where \( r \) is the distance to the origin. Show that \( nr \times \hat{s} \) is a constant along an optical ray traveling in that medium.

**Solution 3.3.4.** The following holds

\[ \frac{d}{ds} (nr \times \hat{s}) = \frac{dr}{ds} \times (n\hat{s}) + r \times \frac{d(n\hat{s})}{ds}. \tag{3.73} \]

The first term on the right hand side of Eq. (3.73) vanishes since \( \frac{dr}{ds} = \hat{s} \) [see Eq. (3.42)]. The same identity (3.42) together with Eq. (3.61) lead to

\[ \frac{d(n\hat{s})}{ds} = \nabla n, \tag{3.74} \]

and thus also the second term on the right hand side of Eq. (3.73) vanishes provided that \( n = n(r) \), which implies that

\[ \frac{d}{ds} (nr \times \hat{s}) = 0, \tag{3.75} \]

i.e. \( nr \times \hat{s} \) is a constant along an optical ray.

### 3.3.3 Fermat’s Principle

In the case of a homogeneous medium the light velocity is given by \( c/n \) [see Eq. (2.153)]. In geometrical optics it is assumed that the length scale characterizing changes in the refraction index \( n \) in the medium is much larger than the optical wavelength, and consequently \( c/n \) can be considered as the local value of light velocity in the medium.

Let \( C \) be a curve given by the parametrization \( x_C(q) \)

\[ x_C(q) = (x_{C1}(q), x_{C2}(q), x_{C3}(q)), \tag{3.76} \]

where \( q \in [q_1, q_2] \). The time of flight \( T \) of light traveling along the curve \( C \) is given by

\[ T = c^{-1} \int_{q_1}^{q_2} dq \, n(x_C) |\dot{x}_C|, \tag{3.77} \]

where

\[ \dot{x}_C = \frac{dx_C}{dq}, \tag{3.78} \]
or, alternatively by
\[ T = c^{-1} \int_{q_1}^{q_2} dq \mathcal{L}(x_C, \dot{x}_C), \]
where \( \mathcal{L} \) is given by
\[ \mathcal{L}(x_C, \dot{x}_C) = n(x_C) \sqrt{\dot{x}_C^2}, \]
and
\[ \dot{x}_C = (\dot{x}_{C1}(q), \dot{x}_{C2}(q), \dot{x}_{C3}(q)). \]

Consider another curve \( C' \), which is assumed to be infinitesimally close to the curve \( C \), and which is given by
\[ x'_C(q) = x_C(q) + \delta x(q). \]
It is assumed that \( \delta x(q_1) = \delta x(q_2) = 0 \), i.e. the curves \( C \) and \( C' \) are taken to have the same initial and final points. To lowest nonvanishing order in \( \delta x(q) = (\delta x_{C1}(q), \delta x_{C2}(q), \delta x_{C3}(q)) \) the change \( \delta \mathcal{L} \) in \( \mathcal{L} \) is given by
\[ \delta \mathcal{L} = \sum_{n=1}^{3} \frac{\partial \mathcal{L}}{\partial x_{Cn}} \delta x_{Cn} + \frac{\partial \mathcal{L}}{\partial \dot{x}_{Cn}} \frac{d}{dq} (\delta C_n), \]
thus, to lowest nonvanishing order in \( \delta x \) the change \( \delta T \) in the time of flight \( T \) is given by
\[ \delta T = c^{-1} \int_{q_1}^{q_2} dq \delta \mathcal{L}(x_C, \dot{x}_C) \]
\[ = c^{-1} \int_{q_1}^{q_2} dq \sum_{n=1}^{3} \frac{\partial \mathcal{L}}{\partial x_{Cn}} \delta x_{Cn} + \frac{\partial \mathcal{L}}{\partial \dot{x}_{Cn}} \frac{d}{dq} (\delta C_n). \]

Integration by parts yields [recall that \( \delta x(q_1) = \delta x(q_2) = 0 \)]
\[ \delta T = c^{-1} \int_{q_1}^{q_2} dq \sum_{n=1}^{3} \left( \frac{\partial \mathcal{L}}{\partial x_{Cn}} - \frac{d}{dq} \frac{\partial \mathcal{L}}{\partial \dot{x}_{Cn}} \right) \delta x_{Cn}, \]
or
\[ \delta T = c^{-1} \int_{q_1}^{q_2} dq \left( \sqrt{n^2 \nabla n - \frac{d}{dq} n \dot{x}_C \sqrt{n^2}} \right) \cdot \delta x. \]

Claim. The requirement the \( \delta T = 0 \) for an arbitrary \( \delta x \) implies that the curve \( C \) satisfies Eq. (3.61).
Proof. The requirement implies that [see Eq. (3.86)]
\[ \frac{d}{dq} n \dot{x}_C = \sqrt{\dot{x}_C^2} \nabla n. \]  
(3.87)

With arc-length parameterization the curve \( C \) is expressed as \( x_C(q(s)) \), where
\[ \left| \frac{dx_C}{ds} \right| = 1. \]  
(3.88)

With the help of the following relation
\[ \frac{d}{dq} = \frac{ds}{dq} \frac{d}{ds}, \]  
(3.89)

Eq. (3.87) becomes
\[ \frac{d}{ds} \left( n \frac{dx_C}{ds} \right) = \nabla n, \]  
(3.90)
i.e. the curve \( C \) satisfies Eq. (3.61).

The above claim demonstrates that the ray equation (3.61) can be obtained by assuming the so-called Fermat’s principle, which states that an optical ray locally minimizes the time of flight of light traveling from a given initial point to a given final point.

**Exercise 3.3.5.** Show that for an optical ray connecting an initial point \( r_1 \) to a final one \( r_2 \) the time of flight \( T \) is given by
\[ T = \frac{\psi(r_2) - \psi(r_1)}{c}, \]  
(3.91)
where \( \psi \) is the eikonal.

**Solution 3.3.5.** With the help of Eqs. (3.42) and (3.77) one finds that for an optical ray \( r(s) \) in arc-length parametrization the following holds
\[ T = c^{-1} \int_{s_1}^{s_2} ds \left| \frac{dr}{ds} \right| n \]  
\[ = c^{-1} \int_{s_1}^{s_2} ds \left( \frac{dr}{ds}, \nabla \psi \right) \]  
\[ = c^{-1} \int_{r_1}^{r_2} dr \cdot \nabla \psi, \]  
(3.92)
thus Eq. (3.91) holds.
Chapter 3. Geometrical Optics

As is demonstrated in the exercise below, transformation of coordinates is commonly made simpler by employing the Fermat’s principle for the derivation of the ray equation.

Exercise 3.3.6. Consider a cylindrically symmetric medium, for which the refractive index \( n \) depends only on the distance \( r = \sqrt{x^2 + y^2} \) to the symmetry axis, which is taken to be the \( z \) axis. Express the ray equation in cylindrical coordinates \((r, \phi, z)\), where \( \phi = \tan^{-1}(y/x) \).

Solution 3.3.6. As was shown above, the ray equation can be obtained from the set of equations [see Eq. (3.85)]

\[
\frac{\partial L}{\partial x_C} = \frac{d}{dq} \frac{\partial L}{\partial \dot{x}_C},
\]

where \( L \) is given by Eq. (3.80). For the case of cylindrical coordinates one obtains

\[
\frac{\partial L}{\partial r} = \frac{d}{dq} \frac{\partial L}{\partial \dot{r}},
\]

\[
\frac{\partial L}{\partial \phi} = \frac{d}{dq} \frac{\partial L}{\partial \dot{\phi}},
\]

\[
\frac{\partial L}{\partial z} = \frac{d}{dq} \frac{\partial L}{\partial \dot{z}}.
\]

The following holds

\[
x_C = (r \cos \phi, r \sin \phi, z),
\]

\[
\dot{x}_C = \left( \frac{dr}{dq} \cos \phi - r \sin \phi \frac{d\phi}{dq}, \frac{dr}{dq} \sin \phi + r \cos \phi \frac{d\phi}{dq}, \frac{dz}{dq} \right),
\]

and thus

\[
x_C^2 = \left( \frac{dr}{dq} \right)^2 + \left( \frac{d\phi}{dq} \right)^2 + \left( \frac{dz}{dq} \right)^2,
\]

and [see Eq. (3.80)]

\[
L = n \sqrt{x_C^2} = n \sqrt{\dot{r}^2 + (r \dot{\phi})^2 + \dot{z}^2}.
\]

With the help of the above result Eqs. (3.94), (3.95) and (3.96) become

\[
\sqrt{x_C^2} \frac{\partial n}{\partial r} = \frac{d}{dq} \frac{n \dot{r}}{\sqrt{x_C^2}},
\]

\[
\sqrt{x_C^2} \frac{\partial n}{\partial \phi} = \frac{d}{dq} \frac{n r^2 \dot{\phi}}{\sqrt{x_C^2}},
\]

\[
\sqrt{x_C^2} \frac{\partial n}{\partial z} = \frac{d}{dq} \frac{n \dot{z}}{\sqrt{x_C^2}}.
\]
With arc-length parameterization the curve $C$ is expressed as $x_C(q(s))$, where
\[
\left| \frac{dx_C}{ds} \right| = 1 , \quad (3.104)
\]
and where
\[
\frac{d}{dq} = \frac{ds}{dq} \frac{d}{ds} , \quad (3.105)
\]
and thus Eqs. (3.101), (3.102) and (3.103) become
\[
\frac{\partial n}{\partial r} = \frac{d}{ds} \left( \frac{n \frac{dr}{ds}}{ds} \right) , \quad (3.106)
\]
\[
\frac{\partial n}{\partial \phi} = \frac{d}{ds} \left( m^2 \frac{d\phi}{ds} \right) , \quad (3.107)
\]
\[
\frac{\partial n}{\partial z} = \frac{d}{ds} \left( \frac{n \frac{dz}{ds}}{ds} \right) . \quad (3.108)
\]
For the case of cylindrically symmetric medium $n = n(r)$, and thus
\[
\frac{\partial n}{\partial r} = \frac{d}{ds} \left( \frac{n \frac{dr}{ds}}{ds} \right) , \quad (3.109)
\]
\[
0 = \frac{d}{ds} \left( m^2 \frac{d\phi}{ds} \right) , \quad (3.110)
\]
\[
0 = \frac{d}{ds} \left( \frac{n \frac{dz}{ds}}{ds} \right) . \quad (3.111)
\]

### 3.4 Transport Equation

The so-called transport equation [see Eq. (3.112) below] is needed in order to evaluate the evolution of polarization along an optical ray.

**Claim.** The vector $E_0$ satisfies the transport equation, which is given by
\[
2 \left( \nabla \psi \cdot \nabla \right) E_0 + E_0 \left[ \nabla^2 \psi - \nabla \left( \log \mu \right) \cdot \nabla \psi \right] + 2 \left[ E_0 \cdot \nabla \left( \log n \right) \right] \nabla \psi = 0 . \quad (3.112)
\]

**Proof.** Eliminating $H_{m+1}$ and $H_m$ [lowering the index in Eq. (3.23) by one] from Eq. (3.23) and substituting into Eq. (3.22) yield
\[
\nabla \times \left[ \frac{1}{\mu} \nabla \times E_{m-1} \right] + \nabla \times \left[ \frac{\nabla \psi}{\mu} \times E_{m} \right] = -\epsilon E_{m+1} - \frac{\nabla \psi}{\mu} \times (\nabla \times E_{m}) - (\nabla \psi) \times \left[ \frac{1}{\mu} (\nabla \psi) \times E_{m+1} \right] , \quad (3.113)
\]
Chapter 3. Geometrical Optics

The two terms involved with $E_m$ are treated as follows. Using the vector identities
\[ \nabla \cdot (A \cdot B) = A \times (\nabla \times B) + B \times (\nabla \times A) + (B \cdot \nabla) A + (A \cdot \nabla) B, \] (3.114)
and (1.127), which is given by
\[ \nabla \times (A \times B) = A (\nabla \cdot B) - B (\nabla \cdot A) + (B \cdot \nabla) A - (A \cdot \nabla) B, \] (3.115)
one finds that
\[ \nabla \times (A \times B) + A \times (\nabla \times B) = A (\nabla \cdot B) - B (\nabla \cdot A) + (B \cdot \nabla) A - (A \cdot \nabla) B \]
(3.116)
thus
\[ \nabla \times \left( \frac{\psi}{\mu} \times E_m \right) + \frac{\nabla \psi}{\mu} \times (\nabla \times E_m) = A \psi \frac{\nabla \psi}{\mu} \cdot E_m - E_m \left( \frac{\nabla \psi}{\mu} \right) \times \left( \nabla \times \frac{\psi}{\mu} \right) - 2 \left( \frac{\nabla \psi}{\mu} \cdot \nabla \right) E_m. \] (3.117)

By using Eq. (2.152) and the vector identity $\nabla \times (\nabla f) = 0$ one obtains
\[ \nabla \times \frac{\nabla \psi}{\mu} = \frac{1}{\mu} \nabla \times \nabla \psi + \nabla \left( \frac{1}{\mu} \right) \times \nabla \psi = \nabla \left( \frac{1}{\mu} \right) \times \nabla \psi. \] (3.118)
Substituting into Eq. (3.113) yields
\[ 2 (\nabla \psi \cdot \nabla) E_m - \nabla \psi (\nabla \cdot E_m) + \mu E_m \left( \frac{\nabla \psi}{\mu} \right) \]
\[ - \mu \nabla \left( \frac{\nabla \psi}{\mu} \cdot E_m \right) + \mu E_m \times \left( \nabla \left( \frac{1}{\mu} \right) \times \nabla \psi \right) = n^2 E_{m+1} + (\nabla \psi) \times [(\nabla \psi) \times E_{m+1}] + \mu \nabla \times \left[ \frac{1}{\mu} \nabla \times E_{m-1} \right], \] (3.119)
where $n = \sqrt{\epsilon \mu}$. By using Eq. (1.95) for the fifth term on the left hand side and for the second term on the right hand side one finds that
\[ 2 (\nabla \psi \cdot \nabla) E_m - \nabla \psi (\nabla \cdot E_m) + \mu E_m \left( \frac{\nabla \psi}{\mu} \right) \]
\[ - \mu \nabla \left( \frac{\nabla \psi}{\mu} \cdot E_m \right) + \mu (E_m \cdot \nabla \psi) \nabla \left( \frac{1}{\mu} \right) - \mu \left( E_m \cdot \nabla \left( \frac{1}{\mu} \right) \right) \nabla \psi \]
\[ = \left[ n^2 - (\nabla \psi)^2 \right] E_{m+1} + [(\nabla \psi) \cdot E_{m+1}] (\nabla \psi) + \mu \nabla \times \left[ \frac{1}{\mu} \nabla \times E_{m-1} \right]. \] (3.120)
With the help of Eq. (3.41) one finds that the first term on the left hand side vanishes. Next, Eq. (3.24), which reads

\[ -\nabla \psi \cdot E_m = \frac{1}{\epsilon} \nabla \cdot (\epsilon E_{m-1}) , \]  

(3.121)
is employed to rewrite the forth term on the left hand side, thus

\[ \nabla \left( \frac{\nabla \psi}{\mu} \cdot E_m \right) = \nabla \left( \frac{1}{\mu} \right) (\nabla \psi \cdot E_m) + \frac{1}{\mu} \nabla (\nabla \psi \cdot E_m) \]
\[ = \nabla \left( \frac{1}{\mu} \right) (\nabla \psi \cdot E_m) - \frac{1}{\mu} \nabla \left[ \frac{1}{\epsilon} \nabla \cdot (\epsilon E_{m-1}) \right] , \]  

(3.122)
hence

\[ 2 (\nabla \psi \cdot \nabla) E_m - \nabla \psi (\nabla \cdot E_m) + \mu E_m \left( \nabla \cdot \frac{\nabla \psi}{\mu} \right) - \mu \left( E_m \cdot \nabla \left( \frac{1}{\mu} \right) \right) \nabla \psi \]
\[ = [(\nabla \psi) \cdot E_{m+1}] (\nabla \psi) + \mu \nabla \times \left[ \frac{1}{\mu} \nabla \times E_{m-1} \right] - \nabla \left[ \frac{1}{\epsilon} \nabla \cdot (\epsilon E_{m-1}) \right] . \]  

(3.123)

Using Eqs. (3.24) and (2.151) for the second term on the left hand side one finds that

\[ -\nabla \psi \nabla \cdot E_m = -\nabla \psi \left[ \frac{1}{\epsilon} \nabla \cdot (\epsilon E_m) - \frac{1}{\epsilon} E_m \cdot \nabla \epsilon \right] \]
\[ = \nabla \psi \left[ \nabla \psi \cdot E_{m+1} + \frac{1}{\epsilon} E_m \cdot \nabla \epsilon \right] , \]  

(3.124)
thus

\[ 2 (\nabla \psi \cdot \nabla) E_m + \nabla \psi \left( \frac{1}{\epsilon} E_m \cdot \nabla \epsilon \right) + \mu E_m \left( \nabla \cdot \frac{\nabla \psi}{\mu} \right) - \mu \left( E_m \cdot \nabla \left( \frac{1}{\mu} \right) \right) \nabla \psi \]
\[ = \mu \nabla \times \left[ \frac{1}{\mu} \nabla \times E_{m-1} \right] - \nabla \left[ \frac{1}{\epsilon} \nabla \cdot (\epsilon E_{m-1}) \right] . \]  

(3.125)

After arranging terms this becomes

\[ 2 (\nabla \psi \cdot \nabla) E_m + E_m \left[ \nabla^2 \psi - \nabla (\log \mu) \cdot \nabla \psi \right] + 2 \left[ E_m \cdot \nabla (\log n) \right] \nabla \psi \]
\[ = \mu \nabla \times \left[ \frac{1}{\mu} \nabla \times E_{m-1} \right] - \nabla \left[ \frac{1}{\epsilon} \nabla \cdot (\epsilon E_{m-1}) \right] . \]  

(3.126)

For the case \( m = 0 \) Eq. (3.126) becomes Eq. (3.112).
3.4.1 Polarization Evolution

In the exercise below the transport equation (3.112) is rewritten in terms of the unit vector $\hat{e}_0$, which points in the direction of $E_0$.

**Exercise 3.4.1.** Show that

$$\frac{d}{ds} \hat{e}_0 = -\kappa (\hat{e}_0 \cdot \hat{\nu}) \hat{s},$$

(3.127)

where $\hat{e}_0$ is a unit vector in the direction of $E_0$

$$\hat{e}_0 \equiv \frac{E_0}{\sqrt{E_0 \cdot E_0^*}}. \quad (3.128)$$

**Solution 3.4.1.** By multiplying the transport equation (3.112) by $E_0^*$, using Eq. (3.36), and taking the real part of the resulting equation one obtains

$$\nabla^2 \psi - \left[ \nabla (\log \mu) \cdot (\nabla\psi) \right] (E_0 \cdot E_0^*) + (\nabla\psi \cdot \nabla) (E_0 \cdot E_0^*) = 0. \quad (3.129)$$

Substituting $E_0 = \sqrt{E_0 \cdot E_0^*} \hat{e}_0$ [see Eq. (3.128)] into Eq. (3.112) leads to

$$\frac{1}{2} \left[ \nabla^2 \psi - \left[ \nabla (\log \mu) \cdot (\nabla\psi) \right] \right] \sqrt{E_0 \cdot E_0^*} \hat{e}_0 + \sqrt{E_0 \cdot E_0^*} (\nabla\psi \cdot \nabla) \hat{e}_0$$

$$+ \frac{\hat{e}_0}{2\sqrt{E_0 \cdot E_0^*}} (\nabla\psi \cdot \nabla) (E_0 \cdot E_0^*) + \left( \sqrt{E_0 \cdot E_0^*} \hat{e}_0 \cdot \nabla (\log n) \right) \nabla \psi = 0,$$

(3.130)

and thus [see Eq. (3.129)]

$$(\nabla\psi \cdot \nabla) \hat{e}_0 + (\hat{e}_0 \cdot \nabla (\log n)) \nabla \psi = 0,$$

(3.131)

or

$$\frac{d}{ds} \hat{e}_0 = -[\hat{e}_0 \cdot \nabla (\log n)] \hat{s},$$

(3.132)

thus Eq. (3.127) holds [see Eq. (3.56)]. Note that with the help of Eq. (1.95), which is given by

$$A \times (B \times C) = (A \cdot C) B - (A \cdot B) C,$$

(3.133)

Eq. (3.127) can be rewritten as [as can be seen from Eq. (3.36), $\hat{e}_0 \cdot \hat{s} = 0$]

$$\frac{d}{ds} \hat{e}_0 = -\kappa \hat{e}_0 \times \hat{b},$$

(3.134)

where $\hat{b} = \hat{s} \times \hat{\nu}$ is the binormal unit vector.
Using the notation
\[ \hat{e}_0 = e_\nu \hat{\nu} + e_b \hat{b}, \] (3.135)
one finds using Eq. (3.44) that
\[ \frac{d e_\nu}{ds} \hat{\nu} + \frac{d e_b}{ds} \hat{b} + e_\nu \left(-\kappa \hat{s} + \tau \hat{b}\right) - e_b \tau \hat{\nu} = -\kappa e_\nu \hat{s}, \]
(3.136)
thus
\[ \frac{d}{ds} \begin{pmatrix} e_\nu \\ e_b \end{pmatrix} = i K_g \begin{pmatrix} e_\nu \\ e_b \end{pmatrix}, \]
(3.137)
where
\[ K_g = \tau \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}. \]
(3.138)
The fact that \( \sigma_2 \) is Hermitian, namely \( \sigma_2^\dagger = \sigma_2 \), ensures that the \( s \) evolution of \( \hat{e}_0 \) is unitary. The solution of Eq. (3.137) reads
\[ \begin{pmatrix} e_\nu (s) \\ e_b (s) \end{pmatrix} = \exp \left( i \sigma_2 \theta \right) \begin{pmatrix} e_\nu (0) \\ e_b (0) \end{pmatrix}, \]
(3.139)
where
\[ \theta = \int_0^s ds' \tau (s'). \]
(3.140)
Using the fact that \( \sigma_2^2 = 1 \), where 1 denotes the \( 2 \times 2 \) identity matrix, one finds that
\[ \exp \left( i \sigma_2 \theta \right) = \sum_{n=0}^{\infty} \frac{(i \sigma_2 \theta)^n}{n!} = \cos \theta + i \sigma_2 \sin \theta, \]
(3.141)
thus
\[ \begin{pmatrix} e_\nu (s) \\ e_b (s) \end{pmatrix} = \begin{pmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{pmatrix} \begin{pmatrix} e_\nu (0) \\ e_b (0) \end{pmatrix}. \]
(3.142)
As is shown by Eq. (3.151) below, for the case of a closed curve, i.e. when \( \hat{s} (s) = \hat{s} (0) \), the following holds
\[ \theta = \Omega, \]
(3.143)
where \( \Omega \) is the solid angle subtends by the closed curve \( \hat{s} (s') \) at the origin.

**Exercise 3.4.2.** Calculate the integrated torsion (3.140) for the case of a closed curve, i.e. when \( \hat{s} (s) = \hat{s} (0) \).
Solution 3.4.2. Consider a family of optical rays $\mathbf{r}(s,u)$, where $s$ is an arc-length parameter along an optical ray for any given value of $u$, i.e. $|d\mathbf{r}/ds| = 1$. Consider the integrated torsion $\tau$ over the following infinitesimal closed curve

$$
\begin{aligned}
d\theta &= \left( \int_{(s-\frac{4\nu}{u}-\frac{4\nu}{u})}^{(s+\frac{4\nu}{u}+\frac{4\nu}{u})} + \int_{(s+\frac{4\nu}{u}-\frac{4\nu}{u})}^{(s+\frac{4\nu}{u}+\frac{4\nu}{u})} + \int_{(s+\frac{4\nu}{u}+\frac{4\nu}{u})}^{(s-\frac{4\nu}{u}+\frac{4\nu}{u})} \right) \tau dsdu.
\end{aligned}
\tag{3.144}
$$

Using the relation

$$
\tau = -\hat{\nu} \cdot \frac{d\hat{b}}{ds} = \hat{b} \cdot \frac{d\hat{\nu}}{ds},
\tag{3.145}
$$

one finds to lowest order in $ds$ and $du$ that

$$
\begin{aligned}
d\theta &= \left[ \hat{b}(s,u) \cdot \left( \frac{d\hat{\nu}(s,u-\frac{du}{2})}{ds} - \frac{d\hat{\nu}(s,u+\frac{du}{2})}{ds} \right) + \hat{b}(s+\frac{du}{2},u) \cdot \left( \frac{d\hat{\nu}(s+\frac{du}{2},u)}{ds} - \frac{d\hat{\nu}(s-\frac{du}{2},u)}{ds} \right) \right] ds
\end{aligned}
\tag{3.146}
$$

In general, for a unit vector $\hat{u}(\alpha)$, where $\alpha$ is a parameter, the following holds

$$
\hat{u} \cdot \frac{d\hat{u}}{d\alpha} = 0,
\tag{3.147}
$$
3.5 Energy Conservation

thus only the components in the $\hat{s}$ direction contribute

$$d\theta = \left[ \left( \hat{s} \cdot \frac{d\hat{b}}{ds} \right) \left( \hat{s} \cdot \frac{d\hat{\nu}}{du} \right) - \left( \hat{s} \cdot \frac{d\hat{b}}{du} \right) \left( \hat{s} \cdot \frac{d\hat{\nu}}{ds} \right) \right] dsdu . \quad (3.148)$$

Moreover, since $\hat{s} \cdot \hat{b} = \hat{s} \cdot \hat{\nu} = 0$, one obtains

$$d\theta = \left[ \left( \frac{d\hat{s}}{ds} \cdot \hat{b} \right) \left( \frac{d\hat{s}}{du} \cdot \hat{\nu} \right) - \left( \frac{d\hat{s}}{du} \cdot \hat{b} \right) \left( \frac{d\hat{s}}{ds} \cdot \hat{\nu} \right) \right] dsdu$$

$$= \hat{s} \cdot \left( \frac{d\hat{s}}{du} \times \frac{d\hat{s}}{ds} \right) dsdu , \quad (3.149)$$

where $dA$ is the area of the infinitesimal closed loop on the unit sphere. This result can be used for evaluating the integral

$$\Delta \theta = \oint_C \tau dsdu \quad (3.150)$$

over a closed loop $C$ by dividing the area into infinitesimally small loops. Thus one concludes that

$$\Delta \theta = \Omega , \quad (3.151)$$

where $\Omega$ is the solid angle that $C$ subtends at the origin.

3.5 Energy Conservation

>From Eq. (3.36) one finds that $E_0$ has no component in the $\hat{s}$ direction, thus one can write

$$E_0 = \alpha_{\nu 0} \hat{\nu} + \alpha_{b 0} \hat{b} . \quad (3.152)$$

Using Eq. (3.35) one finds that

$$H_0 = \sqrt{\frac{\epsilon}{\mu}} \left( \alpha_{\nu 0} \hat{b} - \alpha_{b 0} \hat{\nu} \right) . \quad (3.153)$$

The vectors $E_0$, $H_0$, and $\hat{s}$ are orthogonal to each other. The Poynting phasor vector $S = (c/8\pi) E_0 \times H_0^*$ [see Eqs. (2.95) and (2.139)] is parallel to $\hat{s}$ and real and the following holds

$$\epsilon |E_0|^2 = \mu |H_0|^2 = n (8\pi/c) S \cdot \hat{s} = n (8\pi/c) |S| . \quad (3.154)$$
Chapter 3. Geometrical Optics

The local time averaged electric and magnetic energy densities are given by [see Eq. (2.139)]

\[
\langle w_e \rangle = \frac{\varepsilon}{16\pi} |E_0|^2 ,
\]  
(3.155)

and

\[
\langle w_m \rangle = \frac{\mu}{16\pi} |H_0|^2 ,
\]  
(3.156)

thus in geometrical optics \(\langle w_e \rangle = \langle w_m \rangle\). Denoting the total time averaged energy density as \(\langle w \rangle = \langle w_e \rangle + \langle w_m \rangle\) one finds that

\[
S = v \langle w \rangle \hat{s} ,
\]  
(3.157)

where \(v = c/n\) is the velocity of ray propagation.

As was shown in the previous chapter, energy conservation in a lossless medium leads to a relation between the divergence of the Poynting vector \(S\) and the rate of change in the density of electromagnetic energy \(u\) [see Eq. (2.99)].

**Claim.** In geometrical optics the phasor \(S\) satisfies

\[
\nabla \cdot S = 0 .
\]  
(3.158)

**Proof.** Using Eq. (2.151) one finds that

\[
\mu \nabla \cdot \left( \frac{1}{\mu} \nabla \psi \right) = \nabla^2 \psi - \left[ \nabla (\log \mu) \cdot (\nabla \psi) \right] ,
\]  
(3.159)

thus [see Eq. (3.129)]

\[
\nabla \cdot \left( \frac{\nabla \psi}{\mu} \right) (E_0 \cdot E_0^*) + \left( \frac{\nabla \psi}{\mu} \cdot \nabla \right) (E_0 \cdot E_0^*) = 0 .
\]  
(3.160)

The relations \(\nabla \psi = n \hat{s}\) and \(n = \sqrt{\varepsilon \mu}\) lead to

\[
\nabla \cdot \left( \sqrt{\frac{\varepsilon}{\mu}} \hat{s} \right) (E_0 \cdot E_0^*) + \left( \sqrt{\frac{\varepsilon}{\mu}} \hat{s} \cdot \nabla \right) (E_0 \cdot E_0^*) = 0 ,
\]  
(3.161)

or

\[
\left[ (E_0 \cdot E_0^*) \sqrt{\frac{\varepsilon}{\mu}} \right] \nabla \cdot \hat{s} + \hat{s} \cdot \nabla \left( (E_0 \cdot E_0^*) \sqrt{\frac{\varepsilon}{\mu}} \right) = 0 ,
\]  
(3.162)

thus [see Eq. 2.151]

\[
\nabla \cdot \left[ (E_0 \cdot E_0^*) \sqrt{\frac{\varepsilon}{\mu}} \right] \hat{s} = 0 ,
\]  
(3.163)

in agreement with Eq. (3.158) [see Eq. (3.154)].
3.5. Energy Conservation

3.5.1 Intensity Along an Optical Ray

The intensity $I$ along an optical ray is defined by

$$ I = v \langle w \rangle . \tag{3.164} $$

**Exercise 3.5.1.** Show that

$$ \frac{d}{ds} \log I = -\nabla \cdot \hat{s} . \tag{3.165} $$

**Solution 3.5.1.** With the help of Eqs. (3.157), (3.158) and (3.164) one obtains

$$ 0 = \nabla \cdot \mathbf{S} = \nabla \cdot (I \hat{s}) , \tag{3.166} $$

thus [see Eq. (2.151)]

$$ 0 = I \nabla \cdot \hat{s} + \hat{s} \cdot \nabla I , \tag{3.167} $$

in agreement with Eq. (3.165). Note that with the help of Eq. (3.42) the above result (3.165) can alternatively be written as

$$ \frac{d}{ds} \left( \log \frac{I}{n} \right) = -\frac{1}{n} \nabla^2 \psi . \tag{3.168} $$

As can be seen from Eq. (3.165), the evolution of the intensity $I$ along an optical ray depends on the quantity $\nabla \cdot \hat{s}$. The geometrical meaning of the term $\nabla \cdot \hat{s}$ is discussed below.

Consider a point $P$ on an optical ray. Let $S$ be the surface of constant $\psi$ (eikonal) containing the point $P$. Let $M$ be the plane tangent to $S$ at point $P$ (see Fig. 3.2). Let $x'y'z$ be a coordinate system for which $P$ is the origin, $\hat{z}$ is normal to $M$ and $\hat{x}'$ and $\hat{y}'$ are two orthogonal vectors in $M$. The surface $S$ can be described by a graph

$$(x', y', f(x', y')) , \tag{3.169}$$

where $f(\bar{\rho} = 0) = 0$ and $\bar{\nabla} f(\bar{\rho} = 0) = 0$, and the overbar denotes a vector in $xy$ plane. In general, the Taylor expansion of $f$ around a point $\bar{\rho}$ is given by

$$ f(\bar{\rho} + \delta) = \exp (\delta \cdot \bar{\nabla}) f(\bar{\rho}) $$

$$ = f(\bar{\rho}) + (\delta \cdot \bar{\nabla}) f(\bar{\rho}) + \frac{1}{2!} (\delta \cdot \bar{\nabla})^2 f(\bar{\rho}) + \frac{1}{3!} (\delta \cdot \bar{\nabla})^3 f(\bar{\rho}) + \cdots . \tag{3.170} $$

Thus, to lowest nonvanishing order near $\bar{\rho} = 0$ one has
Chapter 3. Geometrical Optics

Fig. 3.2. The tangent plane.

\[ f(x', y') = \frac{1}{2} \left( x' \frac{\partial}{\partial x'} + y' \frac{\partial}{\partial y'} \right)^2 f \]

\[ = \frac{1}{2} (x', y') \begin{pmatrix} f_{x'x'} & f_{x'y'} \\ f_{y'x'} & f_{y'y'} \end{pmatrix} \begin{pmatrix} x' \\ y' \end{pmatrix} . \]

\[ (3.171) \]

The matrix $F$ is symmetric ($F^T = F$), therefore $F$ has real eigenvalues and $F$ has eigenvectors orthogonal to each other. Thus, it is possible to chose an alternative set of coordinates $xyz$, where as before both $\hat{x}$ and $\hat{y}$ lie in $M$, $\hat{x}$ and $\hat{y}$ are orthogonal to each other and both are orthogonal to $\hat{z}$. The directions $\hat{x}$ and $\hat{y}$ are the principle directions of point $P$ on $S$. With these coordinates to lowest nonvanishing order the following holds

\[ f(x, y) = \frac{1}{2} \left( \frac{x^2}{R_1} + \frac{y^2}{R_2} \right) , \]

\[ (3.172) \]

where the principle curvatures are given by

\[ R_1 = \frac{1}{f_{xx}} , \]

\[ (3.173) \]

and

\[ R_2 = \frac{1}{f_{yy}} . \]

\[ (3.174) \]

Using this coordinate system the eikonal function can be expressed as
\[ \psi(r) = \exp(r \cdot \nabla) \psi(0) \]
\[ = \psi(0) + (r \cdot \nabla) \psi(0) + \frac{1}{2!} (r \cdot \nabla)^2 \psi(0) + \ldots \]
\[ = \psi(0) + x\psi_x + y\psi_y + z\psi_z \]
\[ + \frac{1}{2} (x, y, z) \begin{pmatrix} \psi_{xx} & \psi_{xy} & \psi_{xz} \\ \psi_{yx} & \psi_{yy} & \psi_{yz} \\ \psi_{zx} & \psi_{zy} & \psi_{zz} \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} + \cdots \]

(3.175)

On the surface \( S \) the eikonal function is constant \( \psi = \psi(0) \), thus
\[ \psi(x, y, f(x, y)) = \psi(0) . \]

(3.176)

Substituting this condition in the expansion (3.175) yields
\[ x\psi_x + y\psi_y + \frac{1}{2} \left( \frac{x^2}{R_1} + \frac{y^2}{R_2} \right) \psi_z \]
\[ + \frac{1}{2} \left( x, y, \frac{1}{2} \left( \frac{x^2}{R_1} + \frac{y^2}{R_2} \right) \right) \begin{pmatrix} \psi_{xx} & \psi_{xy} & \psi_{xz} \\ \psi_{yx} & \psi_{yy} & \psi_{yz} \\ \psi_{zx} & \psi_{zy} & \psi_{zz} \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} + \cdots = 0 , \]

thus
\[ \psi_x = \psi_y = 0 , \]

(3.177)

and
\[ \psi_z = -R_1\psi_{xx} = -R_2\psi_{yy} . \]

(3.178)

By using Eq. (2.151) one finds that
\[ \nabla \cdot \hat{s} = \nabla \cdot \frac{\nabla \psi}{\sqrt{\nabla \psi \cdot \nabla \psi}} = \frac{1}{\sqrt{\nabla \psi \cdot \nabla \psi}} \left( \nabla^2 \psi - \frac{1}{2} \frac{\nabla (\nabla \psi \cdot \nabla \psi)}{\nabla \psi \cdot \nabla \psi} \cdot \nabla \psi \right) \]
\[ = \frac{1}{|\psi_z|} \left( \psi_{xx} + \psi_{yy} + \psi_{zz} - \psi_z \right) , \]

(3.179)

thus
\[ \nabla \cdot \hat{s} = -\frac{\psi_z}{|\psi_z|} \left( \frac{1}{R_1} + \frac{1}{R_2} \right) . \]

(3.180)

**Exercise 3.5.2.** Calculate the intensity \( I(s) \) along an optical ray propagating in vacuum.
Solution 3.5.2. With the help of Eq. (3.180) the evolution equation for the intensity \( I(s) \) along an optical ray (3.168) can be expressed in terms of the principle curvatures \( R_1(s) \) and \( R_2(s) \) (it is assumed that \( \psi_z = -|\psi_z| \) and \( n = 1 \); recall that \( \nabla \psi = n \hat{s} \))

\[
\frac{d \log I}{ds} = -\frac{1}{R_1(s)} - \frac{1}{R_2(s)}.
\] (3.181)

As can be seen from the following identity

\[
\frac{d}{ds} \left( \frac{R_1 R_2}{(R_1 + s)(R_2 + s)} \right) = -\frac{1}{R_1 + s} - \frac{1}{R_2 + s},
\] (3.182)

the solution can be taken to be given by

\[
I(s) = \frac{R_1 R_2}{(R_1 + s)(R_2 + s)},
\] (3.183)

where \( R_1 \) and \( R_2 \) are the principle curvatures for \( s = 0 \).

The above result (3.183) for the intensity \( I(s) \) along an optical ray for the case \( n = 1 \) can be used in order to express the electric field \( E(s) \) in geometrical optics along an optical ray [see Eq. (3.20) and note that the relation \( \nabla \psi = n \hat{s} \) for the case of a constant \( n \) implies that along the optical ray \( \psi \) can be taken to be given by \( \psi = ns \)]

\[
E(s) = E(0) e^{ik_0 s} \sqrt{\frac{R_1 R_2}{(R_1 + s)(R_2 + s)}}.
\] (3.184)

3.6 Problems

1. **eikonal approximation in 4-vector formalism** - When the potential 4-vector \( A = (\phi, A_1, A_2, A_3)^T = (\phi, A)^T \) [see Eq. (2.26)] is expressed as

\[
a = A e^{i\Psi},
\] (3.185)

the Maxwell’s equation (2.128) becomes

\[
\partial g \left( \partial^T A e^{i\Psi} \eta - (\partial^T A e^{i\Psi} \eta)^T \right) g = \frac{4\pi}{c} f_{\text{ext}}^T.
\] (3.186)

The left hand side of the equation above (3.186) contains a variety of derivative terms. In the eikonal approximation it is assumed that the terms containing derivatives of \( \Psi \) are much larger than terms containing derivatives of all other variables (i.e. the metric \( g \) (2.114), which depends on the relative permittivity \( \epsilon \), on the relative permeability \( \mu \), and on the velocity 4-vector \( U \), and the envelope 4-vector \( A \)).
3.6. Problems

a) Show that in the eikonal approximation Eq. (3.186) becomes

\[ i \partial_g \left( K^T A^T \eta - \eta A K \right) g = \frac{4\pi}{c} J^T_{\text{ext}} \, , \]  

(3.187)

where

\[ K = \partial \Phi = \left( \frac{1}{c} \frac{\partial \Phi}{\partial t}, \frac{\partial \Phi}{\partial x_1}, \frac{\partial \Phi}{\partial x_2}, \frac{\partial \Phi}{\partial x_3} \right) \, , \]  

(3.188)

b) Show that when the so-called generalized Lorenz gauge condition, which reads

\[ \partial g \eta A = 0 \, , \]  

(3.189)

is imposed Eq. (3.187) becomes

\[ -K g K^T A^T \eta g = \frac{4\pi}{c} J^T_{\text{ext}} \, . \]  

(3.190)

c) Employ to eikonal approximation to show that when the velocity 3-vector \( \mathbf{u} \) vanishes the generalized Lorenz gauge condition (3.189) can be expressed as

\[ 0 = n^2 \frac{\partial \phi}{\partial t} + \nabla \cdot \mathbf{A} = 0 \, . \]  

(3.191)

d) Employ the eikonal approximation to show that when \( \mathbf{u} = 0, J^T_{\text{ext}} = 0 \) and the generalized Lorenz gauge condition is imposed the Maxwell’s equation can be expressed as

\[ 0 = \left( \frac{n^2}{c^2} \frac{\partial^2}{\partial t^2} - \nabla^2 \right) A^T \, . \]  

(3.192)

e) Show that when \( J^T_{\text{ext}} = 0 \) and the generalized Lorenz gauge condition is imposed the Maxwell’s equation (3.190) can be expressed as

\[ 0 = K \eta K^T + \frac{n^2}{c^2} - 1 \left( K \mathbf{U} \right)^2 \, . \]  

(3.193)

f) harmonic time dependency - Consider the case where the phase factor \( \Phi \) has the form

\[ \Phi = -\omega t + k_0 \psi (r) \, , \]  

(3.194)

where the angular frequency \( \omega \) is a constant, \( \psi \) is time independent, and \( k_0 \) is given by

\[ k_0 = \frac{\omega}{c} \, . \]  

(3.195)
For this case Eq. (3.188) becomes

\[ K = k_0 (-1, \nabla \psi) \]  \hspace{1cm} (3.196)

The real function \( \psi(\mathbf{r}) \) is commonly called the eikonal function. Express the relation (3.193) for the case where \( \Psi \) is given by Eq. (3.194) and the generalized Lorenz gauge condition is imposed.

g) Show that when \( \Psi \) is given by Eq. (3.194), the velocity 3-vector \( \mathbf{u} \) vanishes and the generalized Lorenz gauge condition is imposed the following holds

\[ \mathbf{A} \cdot \nabla \psi = n^2 \phi \]  \hspace{1cm} (3.197)

h) Show that in the eikonal approximation when the generalized Lorenz gauge condition is imposed the fields \( \mathbf{E} \) and \( \mathbf{B} \) are given by

\[ \mathbf{E} = i k_0 \mathbf{A}_\perp, \]  \hspace{1cm} (3.198)
\[ \mathbf{B} = \nabla \psi \times \mathbf{E}, \]  \hspace{1cm} (3.199)

where \( \mathbf{A}_\perp \) is the component of \( \mathbf{A} \) perpendicular to \( \nabla \psi \).

2. The components of a given 3-vector \( \mathbf{u} = (u_1, u_2, u_3) \) are assumed to satisfy Eq. (3.8), i.e.

\[ u_n \nabla^2 \psi + 2 (\nabla \psi) \cdot (\nabla u_n) = 0, \]  \hspace{1cm} (3.200)

where \( n \in \{1, 2, 3\} \). The unit vector \( \hat{\mathbf{u}} \) is defined as a normalized vector in the direction of \( \mathbf{u} \), i.e.

\[ \hat{\mathbf{u}} = \frac{\mathbf{u}}{\sqrt{\mathbf{u} \cdot \mathbf{u}^*}} = (\hat{u}_1, \hat{u}_2, \hat{u}_3). \]  \hspace{1cm} (3.201)

a) Show that

\[ 0 = (\nabla \psi) \cdot (\nabla \hat{u}_n). \]  \hspace{1cm} (3.202)

b) Contrary to Eq. (3.202), which is obtained from scalar geometrical optics, Eq. (3.127) for the electric field unit vector \( \hat{\mathbf{e}}_0 \), which is based on the transport equation (3.112), is obtained from a vectorial treatment. Show that these results agree only when the medium is homogeneous.

3. Calculate the torsion \( \tau \) of an helix, which in Cartesian coordinates is given by

\[ \mathbf{r}(s) = (r_x(s), r_y(s), r_z(s)) = (R \cos ks, R \sin ks, \alpha ks), \]  \hspace{1cm} (3.203)

where \( R, k, \) and \( \alpha \) are real constant and \( s \) is a real parameter.

4. The refractive index \( n \) in Cartesian coordinates \( (x_1, x_2, x_3) \) is assumed to depend only on the coordinate \( x_3 \). Consider an optical ray traveling in the plane \( x_2 = 0 \).
3.6. Problems

a) Calculate the trajectory of the ray for the case where the refractive index \( n \) is given by

\[
n = n_0 \exp (-\gamma x_3),
\]

(3.204)

where \( n_0 > 1 \) and \( \gamma > 0 \) are constants. Assume that the ray passes through the origin point \((x_1, x_2, x_3) = (0, 0, 0)\) and the angle between the ray and the \( x_1 \) axis at that point is \( \phi_0 \).

b) Calculate the trajectory of the ray for the case where the refractive index \( n \) is given by

\[
n = n_0 \frac{L}{x_3},
\]

(3.205)

where \( n_0 > 1 \) and \( L > 0 \) are constants. Assume that the ray passes through the point \((x_1, x_2, x_3) = (0, 0, L)\) and the angle between the ray and the \( x_1 \) axis at that point vanishes.

5. Consider a spherically symmetric medium, for which the refractive index \( n \) depends only on the radial coordinate \( r = \sqrt{x^2 + y^2 + z^2} \). Let \( r(s) \) be an arc-length parameterization of an optical ray traveling in the medium. The variables \( L_2 \) and \( L_3 \) are defined by

\[
L_2 = n^2 r^2 \left[ 1 - \left( \frac{dr}{ds} \right)^2 \right],
\]

(3.206)

\[
L_3 = nr^2 \sin^2 \theta \frac{d\phi}{ds},
\]

(3.207)

where, in general, the spherical coordinates \((r, \theta, \phi)\) are related to the Cartesian coordinates \((x, y, z)\) by

\[
(x, y, z) = r (\sin \theta \cos \phi, \sin \theta \sin \phi, \cos \theta).
\]

(3.208)

Calculate the derivatives \( dL_2/ds \) and \( dL_3/ds \) along the optical ray.

6. Consider a spherically symmetric medium, for which the refractive index \( n \) depends only on the radial coordinate \( r = \sqrt{x^2 + y^2 + z^2} \). Calculate the optical rays for the case where

a) \( n(r) \) is given by

\[
n(r) = \sqrt{2 - \left( \frac{r}{R} \right)^2},
\]

(3.209)

where \( R \) is a positive constant (Luneburg lens). Assume that \( r \leq R \).

b) \( n(r) \) is given by

\[
n(r) = \frac{n_0}{1 + \left( \frac{r}{R} \right)^2},
\]

(3.210)

where \( n_0 \) and \( R \) are positive constants (Maxwell’s fish-eye).
7. Eliminating $H_{m+1}$ from Eq. (3.23) and substituting into Eq. (3.22) yield

$$-\mu \nabla \times H_m - \nabla \psi \times (\nabla \times E_m) = n^2 E_{m+1} + (\nabla \psi) \times [(\nabla \psi) \times E_{m+1}] ,$$

(3.211)

thus, for the case $m = 0$ one has

$$-\mu \nabla \times H_0 - \nabla \psi \times (\nabla \times E_0) = n^2 E_1 + (\nabla \psi) \times [(\nabla \psi) \times E_1] .$$

(3.212)

Use the above result (3.212) in order to derive the transport equation (3.112).

8. Show that

$$\hat{\nu} \cdot \nabla \times \hat{s} = 0 ,$$

(3.213)

and

$$\hat{b} \cdot \nabla \times \hat{s} = \kappa .$$

(3.214)

9. Show that the eccentricity $e$ of the polarization ellipse [see Eq. (2.164)] is a constant along an optical ray.

10. Consider a medium whose refractive index $n$ is given in Cartesian coordinates $(x, y, z)$ by

$$n = n_0 \sqrt{1 - K^2 (x^2 + y^2)} ,$$

(3.215)

where both $n_0$ and $K$ are real positive constants. Let $\mathbf{r}(s) = (x(s), y(s), z(s))$ be an arc-length parametrization of an optical ray traveling in the medium, for which the following is assumed to hold

$$x^2(s) + y^2(s) = R^2 ,$$

(3.216)

where $R$ is a positive constants. Let $\hat{e}_0(s)$ be a unit vector pointing in the direction of the electric field. The polarization is assumed to be rectilinear, i.e. the unit vector $\hat{e}_0(s)$ is assumed to be real. Let $\phi_e(s)$ be the angle between $\hat{e}_0(s)$ and the binormal unit vector $\hat{b}$. Calculate $d\phi_e/ds$.

11. Consider a trajectory of a point particle having mass $m$ and charge $q$ from the spacial point $\mathbf{r}_1$ at time $t_1$ to point $\mathbf{r}_2$ at time $t_2$ (as measured in a given inertial frame of reference). The relativistic action $S$ corresponding to a given trajectory is defined by

$$S = -mc^2 \int d\tau - \frac{q}{c} \int d\tau U^T \eta A ,$$

(3.217)

where $\tau$ is the proper time [see Eq. (1.13)], $U = dX/d\tau$ is the velocity 4-vector of the particle [see Eq. (1.63)], $\eta$ is the Minkowski metric (1.14).
3.7 Solutions

and $A = (\phi, A_1, A_2, A_3)^T = (\phi, A)^T$ is the electromagnetic potential 4-vector (2.26). Note that the integral along the trajectory $\int d\tau$ evaluates the time of flight of the particle as measured by a clock that is carried along with the moving particle [compare with Eq. (3.77)]. Note also that, as can be seen from Eq. (1.20), the term $U^T \eta A$ is Lorentz invariant, since

$$
(U')^T \eta A' = U^T A^T \eta A = U^T \eta A.
$$

(3.218)

A trajectory that is consistent with the laws of classical mechanics is called a classical trajectory. The principle of least action states that among all possible trajectories from point $r_1$ at time $t_1$ to point $r_2$ at time $t_2$ the action is locally minimized by a classical trajectory. Employ the principle of least action to derive the classical equations of motion of the particle.

12. Consider a point particle of mass $m$ and charge $q$ moving in an electromagnetic field.

a) Find an equation of motion for the velocity 4-vector $U = dX/d\tau$, where $\tau$ is the proper time.

b) Consider the case where in a frame commoving with the particle the electric $E$ and magnetic $B$ fields are given by $E = E_1 \hat{x}_1$ and $B = B_1 \hat{x}_1$, where both $E_x$ and $B_x$ are constants. Solve the equation of motion for this case.

3.7 Solutions

1. By applying the eikonal approximation to the term $\partial^T A^T e^{i\psi}$ one finds that

$$
\partial^T A^T = \partial^T A e^{i\psi} \simeq i e^{i\psi} K^T A^T = i K^T A^T.
$$

(3.219)

a) With the help of Eq. (3.219) one finds that Eq. (3.186) becomes Eq. (3.187).

b) When generalized Lorenz gauge condition (3.189) is imposed Eq. (3.187) becomes

$$
i \partial g K^T A^T \eta g = \frac{4\pi}{c} \rho_{\text{ext}}^T.
$$

(3.220)

In the eikonal approximation Eq. (3.220) can be replaced by Eq. (3.190) [see Eqs. (3.185), (3.219) and (3.188), and note that it is assumed that terms containing derivatives of $K$ can be neglected as well]. Moreover, in the eikonal approximation Eq. (3.189) can be replaced by

$$
K g \eta A = 0.
$$

(3.221)
c) With the help of the eikonal approximation and Eq. (2.114), which reads \( g = \mu^{-1/2} \left( \eta + \left( \xi/c^2 \right) UU^T \right) \), one finds that Eq. (3.189) can be rewritten as

\[
\partial A + \frac{\xi}{c^2} \partial UU^T \eta A = 0 .
\]  

(3.222)

For a vanishing \( u \) the velocity 4-vector \( U \) becomes \( U = (c, 0, 0, 0)^T \) [see Eq. (2.115)], and thus Eq. (3.191) holds (recall that \( \xi = n^2 - 1 \)).

d) With the help of the eikonal approximation one finds that for this case [see Eqs. (2.114) and (3.219)]

\[
\begin{align*}
\partial gK^T A^T &= \frac{1}{\sqrt{\mu}} \partial \left( \eta + \frac{\xi}{c^2} UU^T \right) \partial T A^T \\
&= \frac{1}{\sqrt{\mu}} \left[ \partial \eta \partial T + \left( n^2 - 1 \right) \frac{1}{c^2} \partial^2 T \right] A^T \\
&= \frac{1}{\sqrt{\mu}} \left( \frac{n^2}{c^2} \partial^2 - \nabla \right) A^T ,
\end{align*}
\]

(3.223)

and thus Eq. (3.220) leads to Eq. (3.192).

e) With the help of Eq. (2.114) one finds that

\[
KgK^T = \frac{1}{\sqrt{\mu}} \left( K\eta K^T + \frac{n^2-1}{c^2} KUU^TK^T \right) ,
\]  

(3.224)

and thus Eq. (3.190) leads to Eq. (3.193).

f) By expressing the velocity 4-vector as \( U = \gamma c (1, \beta)^T \), where \( \beta \) is related to the velocity 3-vector \( u \) by \( \beta = u/c \) [see Eq. (2.115)], one finds that Eq. (3.193) becomes [see Eq. (3.196)]

\[
0 = 1 - (\nabla \psi)^2 + \frac{(n^2 - 1) (1 - \beta \cdot \nabla \psi)^2}{1 - \beta^2} ,
\]  

(3.225)

or

\[
(\nabla \psi)^2 = n^2 + \frac{(n^2 - 1) \left[ (1 - \beta \cdot \nabla \psi)^2 - (1 - \beta^2) \right]}{1 - \beta^2} .
\]  

(3.226)

Note that when \( \beta = 0 \) Eq. (3.226) yields \( (\nabla \psi)^2 = n^2 \).

g) For this case Eq. (3.221) becomes [see Eqs. (2.127) and (3.196)]

\[
K \begin{pmatrix}
  n^2 & 0 & 0 & 0 \\
  0 & 1 & 0 & 0 \\
  0 & 0 & 1 & 0 \\
  0 & 0 & 0 & 1
\end{pmatrix} A = 0 ,
\]  

(3.227)

thus Eq. (3.197) holds.
h) With the help of Eqs. (2.29) and (3.219) one finds that in the eikonal approximation Eq. (2.28) becomes

\[
(iK^T A^T \eta - (iK^T A^T \eta)^T = \begin{pmatrix}
0 & E_1 & E_2 & E_3 \\
-E_1 & 0 & -B_3 & B_2 \\
-E_2 & B_3 & 0 & -B_1 \\
-E_3 & B_2 & B_1 & 0
\end{pmatrix},
\]

(3.228)
or [see Eqs. (2.26) and (3.196)]

\[
(iK^0 \left( \begin{array}{cc}
A_1 - \frac{\partial \phi}{\partial x_1} & 0 \\
0 & A_2 - \frac{\partial \phi}{\partial x_2} \\
0 & A_3 - \frac{\partial \phi}{\partial x_3}
\end{array} \right)
\left( \begin{array}{cc}
A_1 & 0 \\
0 & A_2 \\
0 & A_3
\end{array} \right)
\left( \begin{array}{cc}
\frac{\partial \phi}{\partial x_1} & 0 \\
0 & \frac{\partial \phi}{\partial x_2} \\
0 & \frac{\partial \phi}{\partial x_3}
\end{array} \right)
\right)
\]

\[
\begin{pmatrix}
0 & E_1 & E_2 & E_3 \\
-E_1 & 0 & -B_3 & B_2 \\
-E_2 & B_3 & 0 & -B_1 \\
-E_3 & B_2 & B_1 & 0
\end{pmatrix},
\]

(3.229)

and thus

\[
E = ik_0 (A - \phi \nabla \psi),
\]

(3.230)

and

\[
B = ik_0 (\nabla \psi \times A).
\]

(3.231)

With the help of the gauge condition \(A \cdot \nabla \psi = n^2 \phi\) (3.197) Eq. (3.30) can be rewritten as

\[
E = ik_0 \left(A - \frac{A \cdot \nabla \psi}{n^2} \nabla \psi\right).
\]

(3.232)

According to the eikonal equation (3.7) \((\nabla \psi)^2 = n^2\), and thus \(E = ik_0 A_\perp\), in agreement with Eq. (3.198), where \(A_\perp = A - A_\parallel\) is the component of \(A\) perpendicular to \(\nabla \psi\) and where \(A_\parallel = n^{-2} (A \cdot \nabla \psi) \nabla \psi\) is the component of \(A\) parallel to \(\nabla \psi\). With the help of the above result (3.198) Eq. (3.231) becomes Eq. (3.199).

Thus, both fields \(E\) and \(B\) are perpendicular to \(\nabla \psi\), and they are also perpendicular to each other.

2. With the help of Eq. (3.200) and the relation \(u_n = \sqrt{u \cdot u} \hat{u}_n\) [see Eq. (3.201)] one finds that

\[
0 = \sqrt{u \cdot u} \hat{u}_n \nabla^2 \psi + 2 (\nabla \psi) \cdot (\nabla \left( \sqrt{u \cdot u} \hat{u}_n \right))
= 2 \hat{u}_n \left[ \frac{\sqrt{u \cdot u}}{2} \nabla^2 \psi + (\nabla \psi) \cdot \nabla \left( \sqrt{u \cdot u} \right) \right] + 2 \sqrt{u \cdot u} (\nabla \psi) \cdot (\nabla \hat{u}_n).
\]

(3.233)
Chapter 3. Geometrical Optics

a) The following holds
\[ 2 \sqrt{\mathbf{u} \cdot \mathbf{u}^*} (\nabla \psi) \cdot \nabla \sqrt{\mathbf{u} \cdot \mathbf{u}^*} = (\nabla \psi) \cdot \nabla (\mathbf{u} \cdot \mathbf{u}^*) \]
\[ = \sum_{n=1}^{3} (\nabla \psi) \cdot \nabla (u_n u_n^*) \]
\[ = \sum_{n=1}^{3} (\nabla \psi) \cdot (u_n^* \nabla u_n + u_n \nabla u_n^*) \]  
(3.234)

where [see Eq. (3.200)]
\[ \sum_{n=1}^{3} (\nabla \psi) \cdot (u_n^* \nabla u_n + u_n \nabla u_n^*) = - (\mathbf{u} \cdot \mathbf{u}^*) \nabla^2 \psi , \]  
(3.235)

and thus
\[ (\nabla \psi) \cdot \nabla \sqrt{\mathbf{u} \cdot \mathbf{u}^*} = - \frac{\sqrt{\mathbf{u} \cdot \mathbf{u}^*}}{2} \nabla^2 \psi . \]  
(3.236)

The above results (3.233) and (3.236) yield together to Eq. (3.202).

b) With the help of the relation
\[ \frac{d}{ds} = \frac{\nabla \psi \cdot \nabla}{n} , \]
one finds that Eq. (3.202) can be rewritten as
\[ \frac{d\mathbf{u}}{ds} = 0 , \]  
(3.237)

whereas Eq. (3.127) can be written as [see Eq. (3.132)]
\[ \frac{d\mathbf{e}_0}{ds} = - [\mathbf{e}_0 \cdot \nabla (\log n)] \hat{s} , \]  
(3.238)

thus only when the medium is homogeneous agreement is obtained.

3. The following holds
\[ \frac{dr}{ds} = k (-R \sin ks, R \cos ks, \alpha) . \]  
(3.239)

The condition for \( s \) to be an arc-length parameter reads
\[ 1 = \left| \frac{dr}{ds} \right| = k \sqrt{R^2 + \alpha^2} , \]  
(3.240)

or
\[ k = \frac{1}{\sqrt{R^2 + \alpha^2}} . \]  
(3.241)
The following holds

\[ \hat{s} = \frac{dr}{ds} = \frac{1}{\sqrt{R^2 + \alpha^2}} \left( -R \sin \frac{s}{\sqrt{R^2 + \alpha^2}}, R \cos \frac{s}{\sqrt{R^2 + \alpha^2}}, \alpha \right), \quad (3.242) \]

and

\[ \frac{d\hat{s}}{ds} = \frac{R}{R^2 + \alpha^2} \left( -\cos \frac{s}{\sqrt{R^2 + \alpha^2}} - \sin \frac{s}{\sqrt{R^2 + \alpha^2}}, 0 \right), \quad (3.243) \]

thus

\[ \kappa = \frac{R}{R^2 + \alpha^2}, \quad (3.244) \]

and

\[ \nu = - \left( \cos \frac{s}{\sqrt{R^2 + \alpha^2}}, \sin \frac{s}{\sqrt{R^2 + \alpha^2}}, 0 \right). \quad (3.245) \]

Moreover, one has

\[ \hat{b} = \hat{s} \times \nu = \frac{1}{\sqrt{R^2 + \alpha^2}} \begin{vmatrix} x & y & z \\ R \sin \frac{s}{\sqrt{R^2 + \alpha^2}} & R \cos \frac{s}{\sqrt{R^2 + \alpha^2}} & \alpha \\ -\cos \frac{s}{\sqrt{R^2 + \alpha^2}} - \sin \frac{s}{\sqrt{R^2 + \alpha^2}} & 0 & \end{vmatrix} \]

\[ = \frac{1}{\sqrt{R^2 + \alpha^2}} \left( \alpha \sin \frac{s}{\sqrt{R^2 + \alpha^2}}, -\alpha \cos \frac{s}{\sqrt{R^2 + \alpha^2}}, R \right), \quad (3.246) \]

thus

\[ \frac{d\hat{b}}{ds} = -\frac{\alpha}{R^2 + \alpha^2} \left( -\cos \frac{s}{\sqrt{R^2 + \alpha^2}} - \sin \frac{s}{\sqrt{R^2 + \alpha^2}}, 0 \right) \]

\[ = -\tau \nu, \quad (3.247) \]

where the torsion \( \tau \) is given by

\[ \tau = \frac{\alpha}{R^2 + \alpha^2}. \quad (3.248) \]

4. Let \( \mathbf{x}_C(q) = (x_1(q), x_2(q), x_3(q)) \) be an optical ray traveling through the medium. For the case where the parameter \( q \) is chosen to be the coordinate \( x_1 \) Eqs. (3.80) and (3.93) yield

\[ 0 = \frac{d}{dx_1} \frac{n}{\sqrt{\dot{x}_C^2}}, \quad (3.249) \]

\[ \sqrt{\dot{x}_C^2} \frac{\partial n}{\partial x_3} = \frac{d}{dx_1} \frac{n f}{\sqrt{\dot{x}_C^2}}, \quad (3.250) \]
where
\[ \dot{x}_C^2 = 1 + f^2, \] (3.251)
and where
\[ f = \frac{dx_3}{dx_1}, \] (3.252)
and thus the following holds
\[ \frac{\partial \log n}{\partial x_3} (1 + f^2) = \frac{df}{dx_1}. \] (3.253)

a) For this case \( \partial \log n/\partial x_3 = -\gamma. \) Integrating Eq. (3.253) yields (recall the initial condition \( f = dx_3/dx_1 = \tan \phi_0 \) for \( x_1 = 0 \))
\[ -\gamma x_1 = \tan^{-1} f - \phi_0, \] (3.254)
and thus (recall that \( x_3 = 0 \) for \( x_1 = 0 \))
\[ x_3 = \frac{1}{\gamma} \log \frac{\cos(\gamma x_1 - \phi_0)}{\cos \phi_0}. \] (3.255)

b) For this case \( \partial \log n/\partial x_3 = -1/x_3, \) and thus Eq. (3.253) becomes [see Eq. (3.252)]
\[ -\frac{1}{x_3} \left( 1 + \left( \frac{dx_3}{dx_1} \right)^2 \right) = \frac{d^2x_3}{dx_1^2}. \] (3.256)
The following holds
\[ \frac{d^2x_3^2}{dx_1^2} = \frac{d}{dx_1} \left( 2x_3 \frac{dx_3}{dx_1} \right) = 2 \left( \left( \frac{dx_3}{dx_1} \right)^2 + x_3 \frac{d^2x_3}{dx_1^2} \right), \] (3.257)
and thus Eq. (3.256) can be rewritten as
\[ \frac{d^2x_3^2}{dx_1^2} = -2. \] (3.258)
The initial conditions lead to
\[ x_1^2 + x_3^2 = L^2. \] (3.259)

5. Let \( x_C(q) = (r(q), \theta(q), \phi(q)) \) be a general parametrization of the optical ray. For the case of spherical coordinates the following holds
\[ \dot{x}_C^2 = \dot{r}^2 + \left( r \dot{\theta} \right)^2 + \left( r \sin \theta \dot{\phi} \right)^2, \] (3.260)
where overdot denotes a derivative with respect to the parameter $q$, and thus the ray equations in arc-length parametrization for the case of spherically symmetric medium are given by [compare with Eqs. (3.109), (3.110) and (3.111)]

$$\frac{\partial n}{\partial r} = \frac{d}{ds} \left( n \frac{dr}{ds} \right),$$

(3.261)

$$0 = \frac{d}{ds} \left( nr \frac{d\theta}{ds} \right),$$

(3.262)

$$0 = \frac{d}{ds} \left( nr^2 \sin^2 \theta \frac{d\phi}{ds} \right).$$

(3.263)

The above result (3.263) implies that $dL^2/ds = 0$ [see Eq. (3.207)]. Recall Eq. (3.75), which states that for the case of a spherically symmetric medium the following holds

$$\frac{d}{ds} (nr \times \hat{s}) = 0.$$  

(3.264)

With the help of the general vector identity

$$(A \times B) \cdot (C \times D) = (A \cdot C)(B \cdot D) - (A \cdot D)(B \cdot C),$$

(3.265)

one finds that [see Eqs. (3.42) and (3.206)]

$$(nr \times \hat{s}) \cdot (nr \times \hat{s}) = n^2 r^2 \left( \frac{r}{r} \times \hat{s} \right) \cdot \left( \frac{r}{r} \times \hat{s} \right)$$

$$= n^2 r^2 \left[ 1 - \left( \frac{r}{r} \cdot \frac{dr}{ds} \right)^2 \right]$$

$$= n^2 r^2 \left[ 1 - \left( \frac{dr}{ds} \right)^2 \right]$$

$$= L^2,$$

(3.266)

and thus $dL^2/ds = 0$.

6. In a spherically symmetric medium the vector $nr \times \hat{s}$ is a constant along an optical ray $r(s)$ [see Eq. (3.75)]. Thus, the optical ray is a curve in a plane containing the origin. In spherical coordinates $(r, \theta, \phi)$ the plane is taken to be a plane of a constant angle $\phi$. For this case an optical ray in arc-length parametrization satisfies [see Eq. (3.260)]

$$1 = \left( \frac{dr}{ds} \right)^2 + \left( \frac{d\theta}{ds} \right)^2.$$

(3.267)

The variable $L^2$ given by Eq. (3.206) is a constant along an optical ray in a spherically symmetric medium, and thus
Chapter 3. Geometrical Optics

\[
\frac{dr}{ds} = \sqrt{1 - \frac{L^2}{nr^2} \,}, \tag{3.268}
\]

and [see Eq. (3.267)]

\[
\frac{d\theta}{ds} = \frac{L}{nr^2} \, , \tag{3.269}
\]

and thus

\[
\frac{d\theta}{dr} = \frac{L}{r \sqrt{n^2 r^2 - L^2} \,} \, . \tag{3.270}
\]

or

\[
\frac{d\theta}{d\rho} = \frac{1}{\rho \sqrt{\left(\frac{nr}{R}\right)^2 - 1} \,} \, . \tag{3.271}
\]

where

\[
\rho = \frac{r}{R} \, , \tag{3.272}
\]

\[
l = \frac{L}{R} \, . \tag{3.273}
\]
a) For this case Eq. (3.271) yields [see Eq. (3.209)]

\[
\frac{d\theta}{d\rho} = \frac{1}{\rho \sqrt{\left(\frac{nr}{R}\right)^2 - 1} \,} \, . \tag{3.274}
\]

The variable transformation

\[
\rho = \frac{l}{\sqrt{1 + q}} \, \tag{3.275}
\]

leads to

\[
\frac{d\theta}{dq} = -\frac{1}{2 \sqrt{1 - l^2 - q^2} \,} \, . \tag{3.276}
\]

and the transformation

\[
q = \sqrt{1 - l^2 \cos \psi} \, , \tag{3.277}
\]

to

\[
\frac{d\theta}{d\psi} = \frac{1}{2 \,} \, . \tag{3.278}
\]

and thus
3.7. Solutions

\[ \cos(2\theta + 2\psi_0) = \frac{\left(\frac{L}{R}\right)^2 - 1}{\sqrt{1 - \left(\frac{L}{R}\right)^2}}, \quad (3.279) \]

where \( \psi_0 \) is a constant, which can be taken to be zero. Using the notation

\[ r_0^2 = \frac{L^2}{1 - \sqrt{1 - \left(\frac{L}{R}\right)^2}}, \quad (3.280) \]

\[ \eta = \sqrt{1 - \frac{L^2}{R^2}}, \quad (3.281) \]

Eq. (3.279) can be rewritten as

\[ \left(\frac{r}{r_0}\right)^2 = \frac{1 - \eta}{1 + \eta \cos 2\theta}, \quad (3.282) \]

and thus the optical ray is an ellipse having eccentricity \( e \) given by [compare with Eq. (2.164)]

\[ e = \sqrt{\frac{2\eta}{1 + \eta}} = \frac{2\sqrt{1 - \left(\frac{L}{R}\right)^2}}{1 + \sqrt{1 - \left(\frac{L}{R}\right)^2}}, \quad (3.283) \]

b) For this case Eq. (3.271) yields \[\text{[see Eq. (3.210)]}\]

\[ \frac{d\theta}{d\rho} = \frac{1}{\rho \sqrt{\left(\frac{n_0\rho}{n_1\rho + \rho^2}\right)^2 - 1}}. \quad (3.284) \]

Integration leads to

\[ \theta - \theta_0 = \arcsin \left( \frac{1}{\sqrt{\left(\frac{n_0}{n_1}\right)^2 - 4}} \right), \quad (3.285) \]

where \( \theta_0 \) is a constant, and thus

\[ r \sin (\theta - \theta_0) = \frac{\rho^2 - R^2}{R \sqrt{\left(\frac{n_0 R}{L}\right)^2 - 4}}. \quad (3.286) \]

In Cartesian coordinates

\[ x = r \cos \theta, \quad (3.287) \]

\[ y = r \sin \theta, \quad (3.288) \]

Eq. (3.286) becomes

\[ y \cos \theta_0 - x \sin \theta_0 = \frac{x^2 + y^2 - R^2}{R \sqrt{\left(\frac{n_0 R}{L}\right)^2 - 4}}, \quad (3.289) \]
or
\[ x_0^2 + y_0^2 + R^2 = (x + x_0)^2 + (y - y_0)^2 , \] (3.290)

where
\[ x_0 = R \sqrt{\left( \frac{n_0 R}{2L} \right)^2 - 1 \sin \theta_0} , \] (3.291)
\[ y_0 = R \sqrt{\left( \frac{n_0 R}{2L} \right)^2 - 1 \cos \theta_0} , \] (3.292)
i.e. the optical ray is a circle centered at \((-x_0, y_0)\) having radius
\[ \sqrt{x_0^2 + y_0^2 + R^2} = \frac{n_0 R^2}{2L} . \] (3.293)

7. Using Eqs. (1.95) and (3.41) one finds that
\[ -\mu \nabla \times H_0 - \nabla \psi \times (\nabla \times E_0) = n^2 (\hat{s} \cdot E_1) \hat{s} , \] (3.294)
or [see Eq. (3.35)]
\[ -\mu \left[ \nabla \times \left[ \frac{\nabla \psi}{\mu} \times E_0 \right] + \frac{\nabla \psi}{\mu} \times (\nabla \times E_0) \right] = n^2 (\hat{s} \cdot E_1) \hat{s} . \] (3.295)

Solution existence of the above equation requires that the vector \( F_0 \), which is defined as
\[ F_0 \equiv \nabla \times \left[ \frac{\nabla \psi}{\mu} \times E_0 \right] + \frac{\nabla \psi}{\mu} \times (\nabla \times E_0) , \] (3.296)
is parallel to \( \hat{s} \). In what follows we rewrite \( F_0 \) in a way allowing identifying the components parallel and orthogonal to \( \hat{s} \). Using Eq. (3.116) one finds that
\[ F_0 = \frac{\nabla \psi}{\mu} (\nabla \cdot E_0) - E_0 \left( \nabla \cdot \frac{\nabla \psi}{\mu} \right) + \nabla \left( \frac{\nabla \psi}{\mu} \cdot E_0 \right) - E_0 \times \left( \nabla \times \frac{\nabla \psi}{\mu} \right) - 2 \left( \frac{\nabla \psi}{\mu} \cdot \nabla \right) E_0 . \] (3.297)

The third term on the right vanishes [see Eq. (3.36)]. Moreover, using Eq. (2.152) and \( \nabla \times \nabla \psi = 0 \) one finds that
\[ F_0 = \frac{\nabla \psi}{\mu} (\nabla \cdot E_0) - E_0 \left( \nabla \cdot \frac{\nabla \psi}{\mu} \right) - E_0 \times \left[ \nabla \left( \frac{1}{\mu} \right) \times \nabla \psi \right] - 2 \left( \frac{\nabla \psi}{\mu} \cdot \nabla \right) E_0 . \] (3.298)
In addition, using the vector identity (1.95), which is given by
\[ A \times (B \times C) = (A \cdot C)B - (A \cdot B)C, \] (3.299)
for the third term on the right and using again Eq. (3.36) leads to
\[ \mathbf{F}_0 = \frac{\nabla \psi}{\mu} (\nabla \cdot \mathbf{E}_0) - \mathbf{E}_0 \left( \frac{\nabla \cdot \nabla \psi}{\mu} \right) + \left( \mathbf{E}_0 \cdot \nabla \left( \frac{1}{\mu} \right) \right) \nabla \psi - 2 \left( \frac{\nabla \psi}{\mu} \cdot \nabla \right) \mathbf{E}_0. \] (3.300)
Using Eq. (2.151) for the second term on the right and multiplying by \( \mu \) leads to
\[ \mu \mathbf{F}_0 = \nabla \psi (\nabla \cdot \mathbf{E}_0) - \mathbf{E}_0 \left[ \nabla^2 \psi - \nabla \left( \log \mu \right) \cdot \nabla \psi \right] - \left[ \mathbf{E}_0 \cdot \nabla \left( \log \mu \right) \right] \nabla \psi - 2 \left( \nabla \psi \cdot \nabla \right) \mathbf{E}_0. \] (3.301)
The forth term on the right can be rewritten using Eqs. (3.44) and (3.152) as
\[ (\nabla \psi \cdot \nabla) \mathbf{E}_0 = n \frac{d}{ds} \left( \alpha_{\nu_0} \hat{\nu} + \alpha_{b_0} \hat{b} \right) \]
\[ = n \left[ \frac{d\alpha_{\nu_0}}{ds} \hat{\nu} + \frac{d\alpha_{b_0}}{ds} \hat{b} + \alpha_{\nu_0} \left( -\kappa \hat{s} + \tau \hat{b} \right) - \alpha_{b_0} \tau \hat{\nu} \right] \]
\[ = n \left[ \left( \frac{d\alpha_{\nu_0}}{ds} - \alpha_{b_0} \tau \right) \hat{\nu} + \left( \frac{d\alpha_{b_0}}{ds} + \alpha_{\nu_0} \tau \right) \hat{b} - \alpha_{\nu_0} \kappa \hat{s} \right], \] (3.302)
thus, using Eqs. (3.36) and (3.56) one finds that
\[ \hat{s} \cdot \left[ (\nabla \psi \cdot \nabla) \mathbf{E}_0 \right] = -n \mathbf{E}_0 \cdot \kappa \hat{\nu} = - \mathbf{E}_0 \cdot \nabla n. \] (3.303)
Using the last result together with Eq. (3.301) one can write the condition for \( \mathbf{F}_0 \) to be parallel to \( \hat{s} \) as
\[ 0 = -\mathbf{E}_0 \left[ \nabla^2 \psi - \nabla \left( \log \mu \right) \cdot \nabla \psi \right] - 2 \left( \nabla \psi \cdot \nabla \right) \mathbf{E}_0 - 2 \hat{s} \left( \mathbf{E}_0 \cdot \nabla n \right), \] (3.304)
or
\[ 2 \left( \nabla \psi \cdot \nabla \right) \mathbf{E}_0 + \mathbf{E}_0 \left[ \nabla^2 \psi - \nabla \left( \log \mu \right) \cdot \nabla \psi \right] + 2 \left| \mathbf{E}_0 \cdot \nabla \left( \log n \right) \right| \nabla \psi = 0, \] (3.305)
in agreement with Eq. (3.112).
8. By using Eq. (2.152) and $\psi = n\hat{s}$ one finds that

$$0 = \nabla \times \nabla \psi = \nabla \times (n\hat{s}) = n\hat{s} \times \nabla n + (\nabla n) \times \hat{s},$$

(3.306)

thus, by multiplying by $\hat{s}$ one obtains

$$\hat{s} \cdot \nabla \times \hat{s} = 0,$$

(3.307)

and

$$\nabla \times \hat{s} = \hat{s} \times (\nabla \log n).$$

(3.308)

Therefore, the following hold [see Eq. (3.56)]

$$\hat{\nu} \cdot \nabla \times \hat{s} = \hat{\nu} \cdot [\hat{s} \times (\nabla \log n)] = - (\nabla \log n) \cdot \hat{b} = 0,$$

(3.309)

and

$$\hat{b} \cdot \nabla \times \hat{s} = \hat{b} \cdot [\hat{s} \times (\nabla \log n)] = (\nabla \log n) \cdot \hat{\nu} = \kappa.$$

(3.310)

9. By multiplying Eq. (3.112) by $E_0$ one obtains

$$2E_0 \cdot (\nabla \psi \cdot \nabla) E_0 + E_0^2 \left[ \nabla^2 \psi - \nabla (\log \mu) \cdot \nabla \psi \right] + 2 [E_0 \cdot \nabla (\log n)] (E_0 \cdot \nabla \psi) = 0.$$  

(3.311)

As can be seen from Eq. (3.36), the last term vanishes. Using Eq. (2.151) one finds that

$$\nabla^2 \psi - \nabla (\log \mu) \cdot \nabla \psi = \mu \nabla \left( \frac{\nabla \psi}{\mu} \right),$$

(3.312)

thus

$$2E_0 \cdot \left( \frac{\nabla \psi}{\mu} \cdot \nabla \right) E_0 + E_0^2 \nabla \cdot \left( \frac{\nabla \psi}{\mu} \right) = 0.$$  

(3.313)

Using $\nabla \psi = n\hat{s}$ and $n = \sqrt{\varepsilon \mu}$ lead to

$$2 \sqrt{\frac{\varepsilon}{\mu}} E_0 \cdot (\hat{s} \cdot \nabla) E_0 + E_0^2 \nabla \cdot \left( \sqrt{\frac{\varepsilon}{\mu}} \hat{s} \right) = 0,$$

(3.314)

or

$$\sqrt{\frac{\varepsilon}{\mu}} \hat{s} \cdot \nabla E_0^2 + E_0^2 \nabla \cdot \left( \sqrt{\frac{\varepsilon}{\mu}} \hat{s} \right) = 0,$$

(3.315)

thus [see Eq. (2.151)]

$$\nabla \cdot \left( E_0^2 \sqrt{\frac{\varepsilon}{\mu}} \hat{s} \right) = 0.$$

(3.316)
3.7. Solutions

Define the ratio [see Eq. (2.167)]

\[ \eta = \frac{E_2^2}{|E_0|^2} . \]  \hspace{1cm} (3.317)

According to Eqs. (3.163) and (3.316) both \( |E_0|^2 \) and \( E_0^2 \) have the same \( s \) dependence

\[ E_0^2(s) \sqrt{\frac{\epsilon(s)}{\mu(s)}} = E_0^2(s_0) \sqrt{\frac{\epsilon(s_0)}{\mu(s_0)}} \exp \left[ -\int_{s_0}^{s} ds' (\nabla \cdot \hat{s}) \right] . \] \hspace{1cm} (3.318)

Thus, \( \eta \) is a constant on the ray, and consequently, the eccentricity \( e \) of the polarization ellipse is a constant as well, as can be seen from Eq. (2.164).

10. The ray equation is given by [see Eq. (3.61)]

\[ \frac{d}{ds} \left( n \frac{dr}{ds} \right) = \nabla n^2 = - \frac{n_0^2 K^2 (x\hat{x} + y\hat{y})}{n} . \] \hspace{1cm} (3.319)

Since \( x^2(s) + y^2(s) = R^2 \) the refractive index along the ray is a constant given by

\[ n_r = n_0 \sqrt{1 - K^2 R^2} , \] \hspace{1cm} (3.320)

and thus the ray equation (3.319) becomes

\[ \frac{d^2 r}{ds^2} = -k^2 (x\hat{x} + y\hat{y}) , \] \hspace{1cm} (3.321)

where

\[ k = \frac{K}{\sqrt{1 - K^2 R^2}} . \] \hspace{1cm} (3.322)

The general solution for which \( x^2(s) + y^2(s) = R^2 \) is given by

\[ x(s) = R \cos(ks + \phi_0) , \] \hspace{1cm} (3.323)
\[ y(s) = R \sin(ks + \phi_0) , \] \hspace{1cm} (3.324)
\[ z(s) = z_0 + \alpha ks , \] \hspace{1cm} (3.325)

where \( \phi_0, z_0 \) and \( \alpha \) are constants. The requirement that

\[ 1 = \frac{|dr|}{ds} = \sqrt{k^2 (R^2 + \alpha^2)} , \] \hspace{1cm} (3.326)

yields

\[ \alpha = \frac{\sqrt{1 - K^2 R^2}}{k} . \] \hspace{1cm} (3.327)
As can be seen from Eqs. (3.140) and (3.142)

\[ \frac{d\phi}{ds} = \tau, \]

(3.328)

where \( \tau \) is a torsion, which for the case of an helix is given by Eq. (3.248), thus

\[ \frac{d\phi}{ds} = \frac{\alpha}{R^2 + \alpha^2} = \frac{K\sqrt{1 - 2K^2R^2}}{1 - K^2R^2}, \]

(3.329)

where \( 2K^2R^2 \leq 1 \).

11. The action (3.217) can be expressed in terms of a Lagrangian \( L \) as [see Eq. (1.15)]

\[ S = \int_{t_1}^{t_2} dt \, L, \]

(3.330)

where \( L \) is given by

\[ L = -mc^2\sqrt{1 - \frac{\mathbf{\dot{r}} \cdot \mathbf{\dot{r}}}{c^2}} + \frac{q}{c} \mathbf{\dot{r}} \cdot \mathbf{A} - q\phi, \]

(3.331)

and where overdot denotes a derivative with respect to time \( t \), i.e. \( \mathbf{\dot{r}} = dm/dt = (dx_1/dt, dx_2/dt, dx_3/dt) \). Consider a trajectory

\[ \mathbf{r}(t) = \mathbf{r}_c(t) + \delta \mathbf{r}(t), \]

(3.332)

where \( \mathbf{r}_c(t) \) is assumed to be a classical trajectory and where \( \delta \mathbf{r}(t) \) is considered as infinitesimally small. It is assume that \( \mathbf{r}(t_1) = \mathbf{r}_c(t_1) = \mathbf{r}_1 \) and \( \mathbf{r}(t_2) = \mathbf{r}_c(t_2) = \mathbf{r}_2 \), i.e. \( \delta \mathbf{r}(t_1) = \delta \mathbf{r}(t_2) = 0 \). To lowest order in \( \delta \mathbf{r} = (\delta x_1, \delta x_2, \delta x_3) \) the change in the action \( \delta S \) is given by

\[ \delta S = \int_{t_1}^{t_2} dt \, \delta L = \int_{t_1}^{t_2} dt \sum_{n=1}^{3} \left( \frac{\partial L}{\partial x_n} \delta x_n + \frac{\partial L}{\partial \dot{x}_n} \frac{d}{dt} \delta x_n \right). \]

(3.333)

Integrating the second term by parts leads to

\[ \delta S = \int_{t_1}^{t_2} dt \sum_{n=1}^{3} \left( \frac{\partial L}{\partial x_n} \delta x_n + \frac{d}{dt} \frac{\partial L}{\partial \dot{x}_n} \right) \delta x_n \]

\[ + \sum_{n=1}^{N} \left[ \frac{\partial L}{\partial \dot{x}_n} \delta x_n \right]_{t_1}^{t_2}. \]

(3.334)

The last term vanishes since \( \delta \mathbf{r}(t_1) = \delta \mathbf{r}(t_2) = 0 \). The principle of least action requires that \( \delta S = 0 \) for arbitrary \( \delta x_n \), and thus
\[
\frac{\partial L}{\partial x_n} = \frac{d}{dt} \frac{\partial L}{\partial \dot{x}_n}, \tag{3.335}
\]

The set of equations (3.335) is called the Euler-Lagrange equations. For the coordinate \(x_1\) Eq. (3.335) reads

\[
\frac{d}{dt} \frac{\partial L}{\partial \dot{x}_1} = \frac{\partial L}{\partial x_1}, \tag{3.336}
\]

and the following holds

\[
\frac{d}{dt} \frac{\partial L}{\partial \dot{x}_1} = \mathbf{\dot{p}}_1 + q \left( \frac{\partial A_1}{\partial t} + \dot{x}_1 \frac{\partial A_1}{\partial x_1} + \dot{x}_2 \frac{\partial A_1}{\partial x_2} + \dot{x}_3 \frac{\partial A_1}{\partial x_3} \right), \tag{3.337}
\]

where

\[
p_1 = m \gamma \dot{x}_1, \tag{3.338}
\]

\[
\gamma = \frac{1}{\sqrt{1 - \dot{r}^2}}, \tag{3.339}
\]

and

\[
\frac{\partial L}{\partial x_1} = -q \frac{\partial \phi}{\partial x_1} + q \left( \dot{x}_1 \frac{\partial A_1}{\partial x_1} + \dot{x}_2 \frac{\partial A_2}{\partial x} + \dot{x}_3 \frac{\partial A_3}{\partial x} \right), \tag{3.340}
\]

thus [see Eqs. (2.23) and (2.25)]

\[
\mathbf{\dot{p}}_1 = -q \frac{\partial \phi}{\partial x_1} - q \frac{\partial A_1}{c \partial t} \left[ qE_1 \right. \left. \begin{array}{c}
\dot{x}_2 \left( \frac{\partial A_2}{\partial x_1} - \frac{\partial A_1}{\partial x_2} \right) \\
\left( \nabla \times \mathbf{A} \right)_2 \\
\left( \nabla \times \mathbf{A} \right)_3 \\
\left( \mathbf{\hat{r}} \times \left( \nabla \times \mathbf{A} \right) \right)_1 \\
\end{array} \right], \tag{3.341}
\]

or

\[
\mathbf{\dot{p}}_1 = qE_1 + \frac{q}{c} \left( \mathbf{\hat{r}} \times \mathbf{B} \right)_1. \tag{3.342}
\]

Similar equations are obtained for \(\mathbf{\dot{p}}_2\) and \(\mathbf{\dot{p}}_3\) in the same way. These 3 equations can be written in a 3-vector form as

\[
\mathbf{\dot{p}} = q \left( \mathbf{E} + \frac{1}{c} \mathbf{\hat{r}} \times \mathbf{B} \right), \tag{3.343}
\]
where
\[ p = m\gamma \dot{\mathbf{r}}. \] (3.344)

Alternatively, Eq. (3.343) can be rewritten as [see Eq. (3.344)]
\[ \ddot{\mathbf{r}} = \frac{q}{m\gamma} \left( \mathbf{E} + \frac{1}{c} \dot{\mathbf{r}} \times \mathbf{B} \right) - \frac{\dot{\gamma}}{\gamma} \dot{\mathbf{r}}, \] (3.345)
where [see Eq. (3.339)]
\[ \frac{\dot{\gamma}}{\gamma} = \frac{u}{c} \frac{1 - u^2/c^2}{1 - u^2/c^2}, \] (3.346)
and where
\[ u = |\dot{\mathbf{r}}|. \] (3.347)

With the help of Eqs. (3.343) and (3.344) and the relation
\[ \frac{d\mathbf{p}^2}{dt} = 2\mathbf{p} \cdot \dot{\mathbf{p}}, \] (3.348)
one finds [by multiplying Eq. (3.343) from the left by \( \mathbf{p} \)] that
\[ \frac{1}{2} \frac{d\mathbf{p}^2}{dt} = \mathbf{p} \cdot \dot{\mathbf{p}} = qm\gamma \dot{\mathbf{r}} \cdot \mathbf{E}, \] (3.349)
thus [see Eq. (3.339)]
\[ \dot{u} = q \left( 1 - \frac{u^2}{c^2} \right) \dot{\mathbf{r}} \cdot \mathbf{E}. \] (3.350)

Combining Eqs. (3.345), (3.346) and (3.350) yields
\[ \ddot{\mathbf{r}} = \frac{q}{m\gamma} \left( \mathbf{E} + \frac{1}{c} \dot{\mathbf{r}} \times \mathbf{B} - \frac{\mathbf{E} \cdot \dot{\mathbf{r}}}{c^2} \right). \] (3.351)

12. Recall that the equation of motion in a 3-vector form is given by Eq. (3.343).

a) In general, the force 4-vector \( F = dP/d\tau \) (1.76), where \( P = m\dot{X}/d\tau \) [see Eq. (1.63)], is related to the force 3-vector \( \mathbf{f} = d\mathbf{p}/dt \) (1.78) by [see Eq. (1.80)]
\[ F = \gamma \left( \frac{\mathbf{f} \cdot \dot{\mathbf{r}}}{c}, \mathbf{f} \right)^T, \] (3.352)
where \( \dot{\mathbf{r}} = d\mathbf{r}/dt \) and where \( \gamma = 1/\sqrt{1 - \dot{\mathbf{r}} \cdot \dot{\mathbf{r}}/c^2}. \) For the case of a point particle of charge \( q \) in electromagnetic field the force 3-vector \( \mathbf{f} \) is given by [see Eq. (3.343)]
\[ f = q \left( \mathbf{E} + \frac{1}{c} \hat{\mathbf{r}} \times \mathbf{B} \right), \] (3.353)

and thus with the help of the relation (1.67), which is given by

\[ \frac{1}{d\tau} = \gamma \frac{1}{dt}, \] (3.354)

one finds that

\[ F = \frac{q}{c} \eta \hat{F} \frac{dX}{d\tau}, \] (3.355)

where \( \hat{F} \) is given by [see Eq. (2.29)]

\[
\hat{F} = \begin{pmatrix}
0 & E_1 & E_2 & E_3 \\
-E_1 & 0 & -B_3 & B_2 \\
-E_2 & B_3 & 0 & -B_1 \\
-E_3 & -B_2 & B_1 & 0
\end{pmatrix},
\] (3.356)

and the Minkowski metric \( \eta \) is given by [see Eq. (1.14)]

\[
\eta = \begin{pmatrix}
1 & 0 & 0 & 0 \\
0 & -1 & 0 & 0 \\
0 & 0 & -1 & 0 \\
0 & 0 & 0 & -1
\end{pmatrix},
\] (3.357)

or

\[
\frac{dU}{d\tau} = \frac{q \eta \hat{F}}{mc} U,
\] (3.358)

where

\[ U = \frac{dX}{d\tau}. \] (3.359)

Note that the right hand side of Eq. (3.358) is transformed according to the Lorentz transformation, i.e. [see Eqs. (1.11) and (2.32)]

\[ \Lambda \eta \hat{F} U = \eta \hat{F^\prime} U^\prime. \] (3.360)

b) For the case where \( \mathbf{E} = E_1 \hat{x}_1 \) and \( \mathbf{B} = B_1 \hat{x}_1 \) one has

\[
\frac{q \eta \hat{F}}{mc} = \frac{q}{mc} \begin{pmatrix}
0 & E_1 & 0 & 0 \\
E_1 & 0 & 0 & 0 \\
0 & 0 & 0 & B_1 \\
0 & 0 & -B_1 & 0
\end{pmatrix},
\] (3.361)

and thus [see Eq. (3.358)]
\[ \frac{d}{d\tau} \begin{pmatrix} U_0 \\ U_1 \end{pmatrix} = \frac{qE_1}{mc} \sigma_E \begin{pmatrix} U_0 \\ U_1 \end{pmatrix}, \quad (3.362) \]

\[ \frac{d}{d\tau} \begin{pmatrix} U_2 \\ U_3 \end{pmatrix} = \frac{qB_1}{mc} \sigma_B \begin{pmatrix} U_2 \\ U_3 \end{pmatrix}, \quad (3.363) \]

where the $2 \times 2$ matrices $\sigma_E$ and $\sigma_B$, which are given by

\[ \sigma_E = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad (3.364) \]

\[ \sigma_B = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \quad (3.365) \]

satisfy the relation $\sigma_E^2 = -\sigma_B^2 = 1$, where $1$ is the $2 \times 2$ identity matrix. The solution is thus given by

\[ \begin{pmatrix} U_0(\tau) \\ U_1(\tau) \end{pmatrix} = \exp \left( \frac{qE_1 \tau}{mc} \sigma_E \right) \begin{pmatrix} U_0(0) \\ U_1(0) \end{pmatrix}, \quad (3.366) \]

\[ \begin{pmatrix} U_2(\tau) \\ U_3(\tau) \end{pmatrix} = \exp \left( \frac{qB_1 \tau}{mc} \sigma_B \right) \begin{pmatrix} U_2(0) \\ U_3(0) \end{pmatrix}. \quad (3.367) \]

With the help of the Taylor expansion

\[ \exp (M) = 1 + M + \frac{M^2}{2!} + \frac{M^3}{3!} + \frac{M^4}{4!} + \cdots, \quad (3.368) \]

one obtains

\[ \begin{pmatrix} U_0(\tau) \\ U_1(\tau) \end{pmatrix} = \begin{pmatrix} \cosh \frac{qE_1 \tau}{mc} & \sinh \frac{qE_1 \tau}{mc} \\ \sinh \frac{qE_1 \tau}{mc} & \cosh \frac{qE_1 \tau}{mc} \end{pmatrix} \begin{pmatrix} U_0(0) \\ U_1(0) \end{pmatrix}, \quad (3.369) \]

\[ \begin{pmatrix} U_2(\tau) \\ U_3(\tau) \end{pmatrix} = \begin{pmatrix} \cos \frac{qB_1 \tau}{mc} & \sin \frac{qB_1 \tau}{mc} \\ -\sin \frac{qB_1 \tau}{mc} & \cos \frac{qB_1 \tau}{mc} \end{pmatrix} \begin{pmatrix} U_2(0) \\ U_3(0) \end{pmatrix}. \quad (3.370) \]

As can be seen from the comparison with Eq. (1.186), the particle moves along the $x_1$ axis with a constant proper acceleration given by $qE_1/m$. In the plane spanned by the $x_2$ and $x_3$ axes, on the other hand, the particle undergoes a circular motion having cyclotron frequency given by $qB_1/mc$. 
4. Paraxial Approximation

Consider a medium having cylindrical symmetry along an axis, which is taken to be the $z$ axis. In the paraxial approximation it is assumed that light propagates along the $z$ axis, which is commonly referred to as the optical axis, and it remains confined close to it. This chapter discusses the paraxial approximation for both optical rays and optical waves.

4.1 Paraxial Rays

Consider an optical ray in cylindrical coordinates $(r, \phi, z)$. For simplicity, it is assumed that the angle $\phi$ is kept constant along the ray, i.e. the ray is assumed to be planar. The plane is taken to be the $xz$ plane. In arc-length parametrization the ray $\mathbf{r}(s)$ can be expressed in terms of the angle $\theta(s)$ between the optical ray and the $z$ axis as

$$
\frac{dr}{ds} = (\sin(\theta(s)), 0, \cos(\theta(s))) .
$$

In the paraxial approximation it is assumed that the angle $\theta$ is small.

4.1.1 ABCD Matrix

Let $r(z)$ be the radial coordinate of an optical ray and let $r' = dr/dz$. When the transformation from the input values $r_{in} = r(z_{in})$ and $r'_{in} = r'(z_{in})$ to the output values $r_{out} = r(z_{out})$ and $r'_{out} = r'(z_{out})$ is found to be a linear one, it can be expressed by the so-called ABCD ray matrix

$$
\begin{pmatrix}
  r_{out} \\
  r'_{out}
\end{pmatrix} =
\begin{pmatrix}
  A & B \\
  C & D
\end{pmatrix}
\begin{pmatrix}
  r_{in} \\
  r'_{in}
\end{pmatrix} .
$$

Exercise 4.1.1. Calculate the ABCD ray matrix for the cases of (see Fig. 4.1) (a) translation in a homogeneous medium (b) refraction at a planar interface between a medium having refractive index $n_1$ and a medium having refractive index $n_2$ (c) refraction at a curved interface having radius $R$ between a medium having refractive index $n_1$ and a medium having refractive index $n_2$ (d) transmission through a thin lens having focal length $f$. 
Solution 4.1.1. (a) Optical rays in a homogeneous medium are straight lines, and thus for this case [see Fig. 4.1(a)]
\[
\begin{pmatrix} A & B \\ C & D \end{pmatrix} = \begin{pmatrix} 1 & d \\ 0 & 1 \end{pmatrix} .
\] (4.3)
where \(d = z_{\text{out}} - z_{\text{in}}\). (b) In the paraxial approximation the incidence and transmission angles \(\theta_i\) and \(\theta_t\) are both assumed small, and consequently the Snell's law (3.55), which is given by
\[
n_1 \sin \theta_i = n_2 \sin \theta_t ,
\] (4.4)
can be approximated by the relation
\[
n_1 \theta_i = n_2 \theta_t .
\] (4.5)
In this approximation the ABCD matrix is given by [see Fig. 4.1(b)]
\[
\begin{pmatrix} A & B \\ C & D \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & \frac{n_1}{n_2} \end{pmatrix} .
\] (4.6)
(c) For this case [see Fig. 4.1(c)]
\[
\theta_i \simeq r'_\text{in} + \frac{r_{\text{in}}}{R} ,
\] (4.7)
\[
\theta_t \simeq r'_\text{out} + \frac{r_{\text{in}}}{R} ,
\] (4.8)
and thus [see Eqs. (4.5)]
\[
\begin{pmatrix} A & B \\ C & D \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & \frac{n_1 - n_2}{n_2} \end{pmatrix} .
\] (4.9)
(d) For a lens of focal length \(f\) the focusing condition implies that
\[
\begin{pmatrix} A & B \\ C & D \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & -\frac{1}{f} \end{pmatrix} .
\] (4.10)

Exercise 4.1.2. Calculate the focal distance \(f\) of a lens made of a material having refractive index \(n\). The two surfaces of the lens have radii of curvature \(R_1\) and \(R_2\), respectively. According to the sign convention for radii of curvature, for the case of a biconvex lens, for example, the radius of the surface closest to the light source is taken to be positive, whereas the other radius is taken to be negative.

Solution 4.1.2. The ABCD matrix is given by [see Eqs. (4.9) and (4.10)]
4.1. Paraxial Rays

Fig. 4.1. The ABCD ray matrix.

\[
\begin{pmatrix}
A & B \\
C & D
\end{pmatrix} = \begin{pmatrix}
1 & 0 \\
\frac{1}{n} & 1
\end{pmatrix} \begin{pmatrix}
1 & 0 \\
\frac{1}{n} & 1
\end{pmatrix}
\]

(4.11)

= \begin{pmatrix}
1 & 0 \\
\frac{1}{f} & 1
\end{pmatrix},
\]

(4.12)

which the focal length \( f \) is given by the so-called lensmaker’s equation

\[
\frac{1}{f} = (n - 1) \left( \frac{1}{R_1} - \frac{1}{R_2} \right).
\]

(4.13)

Note that for the cases of a translation (4.3) and a lens (4.10) [see also Eq. (4.50) below] the determinant of the ABCD matrix equals unity and \( A = D \).

Exercise 4.1.3. Let \( M \) be the ABCD matrix of a given optical element, which relates rays entering the element from one interface, which is labelled as \( L \), to rays exiting the element through the opposite interface, which is labelled as \( R \). The element can be positioned in a given point along an optical axis in two orientations. In the first one the interface \( R \) is facing the positive direction of the optical axis and in the other orientation the interface \( R \) is facing the negative direction. What can be said about the matrix \( M \) given that the optical element functions in the same way in both orientations.

Solution 4.1.3. Consider an input and output rays that satisfy the following relation [see Eq. (4.2)]

\[
\begin{pmatrix}
\bar{r}_{\text{out}} \\
\bar{r}_{\text{out}}'
\end{pmatrix} = \begin{pmatrix}
A & B \\
C & D
\end{pmatrix} \begin{pmatrix}
\bar{r}_{\text{in}} \\
\bar{r}_{\text{in}}'
\end{pmatrix}.
\]

(4.15)
Chapter 4. Paraxial Approximation

The inversion symmetry that the optical element is assumed to possess implies that the following must hold (note that when reversing the direction of an optical ray $r \rightarrow r$ and $r' \rightarrow -r'$):

$$
\begin{pmatrix}
  r_{\text{in}} \\
  -r'_{\text{in}}
\end{pmatrix}
= 
\begin{pmatrix}
  A & B \\
  C & D
\end{pmatrix}
\begin{pmatrix}
  r_{\text{out}} \\
  -r'_{\text{out}}
\end{pmatrix},
$$

thus

$$
\begin{pmatrix}
  r_{\text{out}} \\
  r'_{\text{out}}
\end{pmatrix}
= 
\begin{pmatrix}
  1 & 0 \\
  0 & -1
\end{pmatrix}^{-1}
\begin{pmatrix}
  A & B \\
  C & D
\end{pmatrix}^{-1}
\begin{pmatrix}
  1 & 0 \\
  0 & -1
\end{pmatrix}
\begin{pmatrix}
  r_{\text{in}} \\
  r'_{\text{in}}
\end{pmatrix}.
$$

(4.16)

The requirement that both Eqs. (4.15) and (4.17) must hold for any $(r_{\text{in}}, r'_{\text{in}})^T$ implies that

$$
\begin{pmatrix}
  A & B \\
  C & D
\end{pmatrix}
= 
\begin{pmatrix}
  1 & 0 \\
  0 & -1
\end{pmatrix}^{-1}
\begin{pmatrix}
  A & B \\
  C & D
\end{pmatrix}^{-1}
\begin{pmatrix}
  1 & 0 \\
  0 & -1
\end{pmatrix}
\frac{1}{AD - BC}
\begin{pmatrix}
  D & B \\
  C & A
\end{pmatrix},
$$

(4.17)

thus (unless $B = C = 0$)

$$
AD - BC = \det M = 1,
$$

(4.18)

and

$$
A = D.
$$

(4.19)

Below the stability of optical cavities is analyzed using ABCD matrices.

**Exercise 4.1.4.** An optical cavity of length $d$ is formed between two concave and perfectly reflecting mirrors facing each other (both mirrors are centered with respect to the optical axis of the system, and both are normal to the optical axis). The left (right) mirror has a radius of curvature $R_1$ ($R_2$). Under what conditions stable light trapping inside the cavity is possible?

**Solution 4.1.4.** The focal length of a mirror having a radius of curvature $R$ is $R/2$. The ABCD matrix corresponding to an integer number $N$ of cycles back and forth between the two mirrors is given by

$$
\begin{pmatrix}
  A & B \\
  C & D
\end{pmatrix}
= M_0^N,
$$

(4.20)

where
4.1. Paraxial Rays

\[ M_0 = M_2 M_1, \]  
(4.22)

and where \( M_1 \) and \( M_2 \) are given by [see Eqs. (4.3) and (4.10)]

\[ M_n = \begin{pmatrix} 1 & 0 \\ -\frac{2}{R_n} & 1 \end{pmatrix} \begin{pmatrix} 1 & d \\ 0 & 1 \end{pmatrix}, \]  
(4.23)

The following holds \( \det(M_n) = 1 \), and thus \( \det(M_0) = 1 \), and therefore the eigenvalues \( \lambda_{\pm} \) of \( M_0 \) are given by

\[ \lambda_{\pm} = \frac{\Tr(M_0)}{2} \pm \sqrt{\left(\frac{\Tr(M_0)}{2}\right)^2 - 1}. \]  
(4.24)

Light trapping inside the cavity is expected to be stable when the absolute values of both eigenvalues of \( M_0 \) do not exceed unity, a condition that is satisfied provided that

\[ \left| \frac{\Tr(M_0)}{2} \right| \leq 1. \]  
(4.25)

The trace of \( M_0 \) is given by

\[ \frac{\Tr(M_0)}{2} = 2 - \frac{4d}{R_1} - \frac{4d}{R_2} + \frac{4d^2}{R_1 R_2}, \]  
(4.26)

and thus the condition (4.25) can be expressed as

\[ 0 \leq \frac{\Tr(M_0) + 2}{4} = g_1 g_2 \leq 1, \]  
(4.27)

where

\[ g_n = 1 - \frac{d}{R_n}. \]  
(4.28)

4.1.2 Möbius Transformation

The intersection points \( z_{\text{in}} \) and \( z_{\text{out}} \) of the (extrapolated) input and output rays, respectively, with the optical axis are given by

\[ z_{\text{in}} = \frac{r_{\text{in}}}{r'_{\text{in}}}, \]  
(4.29)

\[ z_{\text{out}} = \frac{r_{\text{out}}}{r'_{\text{out}}}. \]  
(4.30)

The following holds

\[ z_{\text{out}} = \frac{r_{\text{out}}}{r'_{\text{out}}} = \frac{A r_{\text{in}} + B r'_{\text{in}}}{C r_{\text{in}} + D r'_{\text{in}}} = f_{A,B,C,D}(z_{\text{in}}), \]  
(4.31)
where
\[ f_{A,B,C,D}(z_{in}) = \frac{Az_{in} + B}{Cz_{in} + D}. \] (4.32)

Note that \( f_{A,B,C,D}(z_{in}) \) is a Möbius transformation.

**Claim.** The following holds
\[ f_{A_2,B_2,C_2,D_2}(f_{A_1,B_1,C_1,D_1}(z)) = f_{A,B,C,D}(z), \] (4.33)
where
\[ \left( \begin{array}{cc} A & B \\ C & D \end{array} \right) = \left( \begin{array}{cc} A_2 & B_2 \\ C_2 & D_2 \end{array} \right) \left( \begin{array}{cc} A_1 & B_1 \\ C_1 & D_1 \end{array} \right). \] (4.34)

**Proof.** With the help of Eq. (4.32) one finds that
\[ f_{A_2,B_2,C_2,D_2}(f_{A_1,B_1,C_1,D_1}(z)) = \frac{A_2Az_{in} + B_2}{C_2Az_{in} + D_2} + B_2, \]
\[ = f_{A,B,C,D}(z) \]
\[ = (A_2A_1+C_2B_1)z + A_2B_1 + B_2D_1 \]
\[ = (C_2A_1 + D_2C_1)z + C_2B_1 + D_2D_1, \] (4.35)

thus Eq. (4.33) holds.

### 4.1.3 Ray Equation

For the case of cylindrically symmetric medium the refractive index \( n \) depends on \( r \) only. Recall that for this case the ray equations in arc-length parametrization are given by Eqs. (3.109), (3.110) and (3.111).

**Exercise 4.1.5.** Show that
\[ \frac{d^2r}{dz^2} = \frac{1}{2n_c^2} \frac{dn^2}{dr}, \] (4.36)

where \( n_c \) is a constant.

**Solution 4.1.5.** With the help of Eq. (3.111), which is given by
\[ 0 = \frac{d}{ds} \left( n \frac{dz}{ds} \right), \] (4.37)

one finds that
\[ n \frac{dz}{ds} = n_c, \] (4.38)
where \( n_c \) is a constant. By employing the relation [see Eq. (4.38)]

\[
\frac{d}{ds} = \frac{dz \ d n_c}{dz} = \frac{n_c}{n} \frac{d}{dz},
\]

(4.39)

Eq. (3.109), which is given by

\[
\frac{dn}{dr} = \frac{d}{ds} \left( n \frac{dr}{ds} \right),
\]

(4.40)

becomes

\[
\frac{dn}{dr} = \frac{n_c^2 d^2 r}{n \ dz^2},
\]

(4.41)

or

\[
\frac{1}{2n_c^2} \frac{dn^2}{dr} = \frac{d^2 r}{dz^2}.
\]

(4.42)

**Exercise 4.1.6.** Consider a planar optical ray, for which \( \phi \) is assumed to be a constant. Show that

\[
\left( \frac{dr}{dz} \right)^2 = \frac{n^2 - n_c^2}{n_c^2},
\]

(4.43)

where \( n_c \) is a constant.

**Solution 4.1.6.** When \( \phi \) is a constant the requirement that \( |dr/ds| = 1 \) implies that [see Eq. (3.99)]

\[
1 = \left( \frac{dr}{ds} \right)^2 + \left( \frac{dz}{ds} \right)^2.
\]

(4.44)

Multiplying Eq. (4.36) by \( dr/dz \) yields

\[
\frac{1}{2n_c^2} \frac{dn^2}{dz} = \frac{dr d^2 r}{dz dz^2},
\]

(4.45)

or

\[
\frac{d}{dz} \left( \frac{dr}{dz} \right)^2 - \frac{n^2}{n_c^2} = 0,
\]

(4.46)

thus

\[
\left( \frac{dr}{dz} \right)^2 - \frac{n^2}{n_c^2} = C,
\]

(4.47)

where \( C \) is a constant. On the other hand [see Eqs. (4.38) and (4.44)]
\[
\left( \frac{dr}{dz} \right)^2 - \frac{n^2}{n_c^2} = \left( \frac{dr}{dz} \right)^2 - \frac{n^2}{n_c^2} = 1 - \frac{n^2}{n} = -1 ,
\]

and thus Eq. (4.43) holds.

### 4.1.4 Graded Index Medium

The index of refraction in a graded index (GRIN) medium is given by

\[
n_{\text{GRIN}}(r) = n_0 \sqrt{1 - g^2 r^2} ,
\]

where both \( n_0 \) and \( g \) are constants.

**Claim.** The ABCD ray matrix of a GRIN medium is given by

\[
\begin{pmatrix}
A & B \\
C & D
\end{pmatrix} = \begin{pmatrix}
\cos (gz) & \frac{1}{g} \sin (gz) \\
-g \sin (gz) & \cos (gz)
\end{pmatrix} .
\]

**Proof.** With the help of Eqs. (4.36) and (4.49) one finds that

\[
\frac{d^2 r}{dz^2} = -n_0^2 g^2 r .
\]

Recall that the constant \( n_c \) is given by \( n_c = n \cos \theta \) [see Eqs. (4.1) and (4.38)].

In the paraxial approximation it is assumed that \( \theta \ll 1 \) and thus the equation of motion becomes

\[
\frac{d^2 r}{dz^2} = -g^2 r .
\]

The solution thus reads

\[
r (z) = r_0 \cos (gz) + \frac{r'_0}{g} \sin (gz) ,
\]

where both \( r_0 \) and \( r'_0 \) are constants. The derivative \( r' = dr/dz \) is given by

\[
r' (z) = -g r_0 \sin (gz) + r'_0 \cos (gz) .
\]

The above results (4.53) and (4.54) lead to Eq. (4.50) [see Eq. (4.2)].

As can be seen from Eq. (4.50) the ABCD matrix is periodic in \( z \), and the period is given by the so-called pitch \( p \) of the medium, which is given by

\[
p = \frac{2\pi}{g} .
\]
4.2 Paraxial Waves

Recall that in the scalar approximation all three components of the vector fields $\mathbf{E}$ and $\mathbf{H}$ satisfy the Helmholtz equation (2.153), which is given by

\[
(\nabla^2 + n^2 k_0^2) \psi = 0 ,
\]

(4.56)

where $n$ is the refractive index and where [see Eq. (2.144)]

\[
k_0 = \frac{\omega}{c} .
\]

(4.57)

4.2.1 Paraxial Approximation

Consider a solution having the form

\[
\psi = A(x, y, z) e^{i n_a k_0 z} ,
\]

(4.58)

where the constant $n_a$ represents a characteristic value of the refractive index $n$ in the medium. Substituting into Eq. (4.56) yields

\[
\nabla_{\perp}^2 A + \frac{\partial}{\partial z} \left( \frac{\partial A}{\partial z} + 2in_a k_0 A \right) + (n^2 - n_a^2) k_0^2 A = 0 ,
\]

(4.59)

where

\[
\nabla_{\perp}^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} .
\]

(4.60)

In the paraxial approximation, which assumes

\[
\left| \frac{\partial A}{\partial z} \right| \ll 2n_a k_0 |A| ,
\]

this becomes

\[
i \frac{\partial A}{\partial z} = \mathcal{H} A ,
\]

(4.61)

where

\[
\mathcal{H} = -\frac{1}{2n_a k_0} \nabla_{\perp}^2 - \frac{(n^2 - n_a^2) k_0^2}{2n_a} .
\]

(4.62)

4.2.2 Gaussian Beam in GRIN Medium

Consider a GRIN medium, whose refractive index $n_{\text{GRIN}}$ is given by Eq. (4.49). For this case Eq. (4.61) becomes (the constant $n_a$ is chosen to be given by $n_a = n_0$)
\[
\frac{i}{\partial z} A = \left( -\frac{1}{2n_0k_0} \nabla_\perp^2 + \frac{n_0k_0g^2r^2}{2} \right) A. \tag{4.63}
\]

Consider a gaussian beam solution, for which \( A \) has the form

\[
A = A_0 e^{-i \left( \rho(z) - \frac{n_0k_0r^2}{2q(z)} \right) / 2}, \tag{4.64}
\]

where \( A_0 \) is a constant. The complex beam parameter \( q(z) \) is expressed as

\[
\frac{1}{q(z)} = \frac{1}{R(z)} + \frac{2i}{n_0k_0w^2(z)}, \tag{4.65}
\]

where \( R(z) \) is the radius of curvature of the Gaussian beam and \( w(z) \) is the spot size. The following claim demonstrates the so-called ABCD law.

**Claim.** The complex beam parameter evolves along the \( z \) axis according to

\[
q(z) = Aq_0 + B \frac{Cq_0 + D}{Cq_0 + D}, \tag{4.66}
\]

where \( q_0 = q(z = 0) \) and the \( A, B, C \) and \( D \) parameters are the elements of the ABCD ray matrix that characterizes the propagation of optical rays in the medium [see Eq. (4.50)]

\[
\begin{pmatrix}
A & B \\
C & D
\end{pmatrix} = \begin{pmatrix}
\cos (gz) & \frac{1}{g} \sin (gz) \\
-g \sin (gz) & \cos (gz)
\end{pmatrix}. \tag{4.67}
\]

**Proof.** In cylindrical coordinates \((r, \phi, z)\) the following holds

\[
\nabla_\perp^2 = \frac{1}{r} \frac{\partial}{\partial r} r \frac{\partial}{\partial r} + \frac{1}{r^2} \frac{\partial^2}{\partial \phi^2}, \tag{4.68}
\]

and thus Eq. (4.63) yields

\[
\frac{dP}{dz} + \frac{i}{q} + \frac{n_0k_0r^2}{2q^2} \left( \frac{dq}{dz} - 1 - g^2q^2 \right) = 0. \tag{4.69}
\]

The above must hold for every \( r \), and thus

\[
\frac{dq}{dz} = 1 + g^2q^2, \tag{4.70}
\]

and

\[
\frac{dP}{dz} = -\frac{i}{q}. \tag{4.71}
\]

The general solution of Eq. (4.70) reads
4.2 Paraxial Waves

\[ q(z) = \frac{\tan (g z + C)}{g} , \tag{4.72} \]

where \( C \) is a constant. In terms of the initial value \( q_0 = q(z = 0) \), which is related to \( C \) by

\[ q_0 = \frac{\tan C}{g} , \tag{4.73} \]

one finds with the help of the identity

\[ \tan (x + y) = \frac{\tan x + \tan y}{1 - \tan x \tan y} , \tag{4.74} \]

that

\[ q(z) = \frac{\cos (g z) q_0 + \frac{1}{g} \sin (g z)}{-g \sin (g z) q_0 + \cos (g z)} , \tag{4.75} \]

in agreement with Eq. (4.66).

4.2.3 Homogeneous Case

In the homogeneous limit, i.e. in the limit \( g \to 0 \), Eq. (4.75) becomes [see Eq. (4.70)]

\[ \lim_{g \to 0} q(z) = q_0 + z . \tag{4.76} \]

Consider the case where the beam’s waist is located at \( z = 0 \). For that case \( 1/R(z = 0) = \frac{1}{2} \) [see Eq. (4.65)]. The width of the waist is denoted by \( w_0 \), i.e. [see Eq. (4.65)]

\[ \frac{1}{q_0} = \frac{2i}{n_0 k_0 w_0^2} . \tag{4.77} \]

With the help of Eq. (4.76) one finds that [see Eq. (4.65)]

\[ R(z) = z \left( 1 + \frac{|q_0|^2}{z^2} \right) , \tag{4.78} \]

and

\[ w^2(z) = w_0^2 \left( 1 + \frac{z^2}{|q_0|^2} \right) . \tag{4.79} \]

As can be seen from Eqs. (4.65), (4.78) and (4.79), far from the beam’s waist (i.e. when \(|z| \gg |q_0|\)) the following holds.
\[ \frac{1}{q(z)} \simeq \frac{1}{R(z)} \simeq \frac{1}{z}. \]  \hfill (4.80)

This result suggests that far from the beam’s waist the gaussian beam parameter \( q(z) \) coincides with the intersection point of ray optics. Recall that the same ABCD Möbius transformation is employed for (a) relating the intersection point \( z_{\text{out}} \) of an output ray to the the intersection point \( z_{\text{in}} \) of an input ray [see Eq. (4.32)], and (b) relating an output gaussian beam parameter \( q(z) \) to an input gaussian beam parameter \( q_0 \) [see Eq. (4.66)]. This observation is consistent with the ABCD law, according to which the same coefficients \( A, B, C \) and \( D \) are employed in both cases (for a given cylindrically symmetric optical system).

**Claim.** The ABCD law is applicable for the case of translation in a homogeneous medium.

**Proof.** The claim is easily proved with the help of Eq. (4.76), according to which

\[ q(z) = \frac{Aq_0 + B}{Cq_0 + D}, \]  \hfill (4.81)

where the parameters \( A, B, C \) and \( D \) are the elements of the ABCD ray matrix (4.3) that characterizes the propagation of optical rays in a homogeneous medium, which is given by

\[ \begin{pmatrix} A & B \\ C & D \end{pmatrix} = \begin{pmatrix} 1 & d \\ 0 & 1 \end{pmatrix}, \]  \hfill (4.82)

where \( d \) is the translation distance along the optical axis.

### 4.3 Fiber Bragg Grating

Consider an optical fiber extended along the \( z \) axis and having refractive index given by

\[ n(x, y, z) = \sqrt{n_0^2(x, y) + n_p^2}, \]  \hfill (4.83)

where the term \( n_p = n_p(x, y, z) \) is considered as a perturbation. In the scalar approximation the field components are required to satisfy the Helmholtz equation (4.56)

\[ \left( \nabla^2_\perp + \frac{d^2}{dz^2} + k_0^2n^2 \right) \psi = 0, \]  \hfill (4.84)

where \( \nabla^2_\perp \) is given by Eq. (4.60), \( k_0 = \omega/c \), \( \omega \) is the angular frequency of optical field and \( c \) is light velocity in vacuum. Near the frequency of interest \( \omega \simeq \omega_0 \) solutions of the unperturbed problem, for which \( n_p = 0 \), are given by
\[ \psi_\pm(x, y, z) = \psi_0(x, y) e^{\pm i \beta z}. \]  
\hspace{1cm} (4.85)

The dispersion relation \( \beta(\omega) \) in that region is assumed to be linear and approximately given by
\[
\beta(\omega) = \frac{\omega n_{\text{eff}}}{c} = k_0 n_{\text{eff}},
\hspace{1cm} (4.86)
\]
where \( n_{\text{eff}} \) is the mode effective refractive index (near \( \omega_0 \)). This assumption implies, as can be seen from Eq. (4.84), that the function \( \psi_0(x, y) \) satisfies the following transverse equation
\[
\left\{ \nabla_\perp^2 + k_0^2 \left[ n_0^2(x, y) - n_{\text{eff}}^2 \right] \right\} \psi_0 = 0.
\hspace{1cm} (4.87)
\]

Consider a solution to the perturbed problem having the form
\[
\psi(x, y, z) = \psi_0(x, y) \left[ A_+(z) e^{i \beta z} + A_-(z) e^{-i \beta z} \right].
\hspace{1cm} (4.88)
\]
Substituting into Eq. (4.84) and employing Eq. (4.87) yields
\[
\frac{d^2}{dz^2} \left[ \frac{1}{k_0^2} \left( n_p^2 + n_{\text{eff}}^2 \right) \right] \psi_0 \left( A_+ e^{i \beta z} + A_- e^{-i \beta z} \right) = 0.
\hspace{1cm} (4.89)
\]
The envelope functions \( A_\pm(z) \), which become constants in the unperturbed case, are assumed to be nearly constant on the length scale of a single wavelength, and therefore
\[
\frac{d^2}{dz^2} \left( A_\pm e^{\pm i \beta z} \right) \approx \left( \pm 2 i \beta \frac{dA_\pm}{dz} - \beta^2 A_\pm \right) e^{\pm i \beta z}.
\hspace{1cm} (4.90)
\]
Employing this approximation in Eq. (4.89) leads to
\[
2i \beta \left( \frac{dA_+}{dz} e^{i \beta z} - \frac{dA_-}{dz} e^{-i \beta z} \right) \psi_0
+ \left[ k_0^2 \left( n_p^2 + n_{\text{eff}}^2 \right) - \beta^2 \right] \psi_0 \left( A_+ e^{i \beta z} + A_- e^{-i \beta z} \right) = 0.
\hspace{1cm} (4.91)
\]
Multiplying by \( \psi_0^* \) and integrating over the xy plane yield
\[
2i \beta \left( \frac{dA_+}{dz} e^{i \beta z} - \frac{dA_-}{dz} e^{-i \beta z} \right)
+ \left[ k_0^2 n_{\text{eff}}^2 (1 + 2D) - \beta^2 \right] \left( A_+ e^{i \beta z} + A_- e^{-i \beta z} \right) = 0,
\hspace{1cm} (4.92)
\]
where
\[
D(z) = \frac{\int \int dx dy |\psi_0|^2 n_p^2}{\int \int dx dy |\psi_0|^2}.
\hspace{1cm} (4.93)
\]
Chapter 4. Paraxial Approximation

The coupling term $D(z)$ for the case of Bragg grating is assumed to have the form

$$
D(z) = D(z) e^{\frac{2\pi i}{\Lambda}} + D^*(z) e^{-\frac{2\pi i}{\Lambda}},
$$

(4.94)

where the envelope function $D(z)$ is assumed to be nearly constant on the length scale of the grating period $\Lambda$. The optical angular frequency $\omega$ is assumed to be close to the Bragg frequency $\omega_B$ (i.e. $\omega_0$ is taken to be equal to $\omega_B$), which is given by

$$
\omega_B = \frac{\pi c}{\Lambda n_{\text{eff}}} = \frac{2\pi}{\lambda_B},
$$

(4.95)

were $\lambda_B = 2n_{\text{eff}}\Lambda$ is the Bragg wavelength, i.e. $|\omega - \omega_B| \ll \omega_B$. Thus the factor $\beta$ can be expressed in terms of the normalized detuning factor $\delta$ as

$$
\beta = \frac{\pi}{\Lambda} (1 - \delta),
$$

(4.96)

where

$$
\delta = -\frac{\omega - \omega_B}{\omega_B} \simeq \frac{\Delta \lambda}{\lambda_B} \ll 1,
$$

(4.97)

and where $\Delta \lambda = \lambda - \lambda_B$ is the offset wavelength. To first order in $\delta$

$$
k_0^2 n_{\text{eff}}^2 - \beta^2 \simeq 2 \left( \frac{\pi}{\Lambda} \right)^2 \delta.
$$

(4.98)

With these notations and approximations one obtains

$$
\left( \frac{dA_+}{dz} - \frac{\pi i \delta}{\Lambda} A_+ \right) e^{i\beta z} - \left( \frac{dA_-}{dz} + \frac{\pi i \delta}{\Lambda} A_- \right) e^{-i\beta z} - \frac{\pi i}{\Lambda} D (A_+ e^{i\beta z} + A_- e^{-i\beta z}) = 0.
$$

(4.99)

Collecting terms oscillating close to $\exp(\pi iz/\Lambda)$ yields

$$
\frac{dA_+}{dz} - \frac{\pi i \delta}{\Lambda} A_+ - \frac{\pi i}{\Lambda} D e^{2\pi i z/\Lambda} A_- = 0,
$$

(4.100)

whereas collecting terms oscillating close to $\exp(-\pi iz/\Lambda)$ yields

$$
\frac{dA_-}{dz} + \frac{\pi i \delta}{\Lambda} A_- + \frac{\pi i}{\Lambda} D^* e^{-2\pi i z/\Lambda} A_+ = 0.
$$

(4.101)

In order to ensure that small terms are kept only up to first order (recall that both $D$ and $\delta$ are considered to be small) the terms $\exp(\pm2\pi i \delta z/\Lambda)$ are replaced by unity. In terms of the dimensionless displacement $\zeta = \pi z/\Lambda$ these two coupled equations can be expressed as
4.3. Fiber Bragg Grating

\[
\begin{align*}
\frac{dA_+}{d\zeta} - i\delta A_+ - iDA_- &= 0, \\
\frac{dA_-}{d\zeta} + i\delta A_- + iD^*A_+ &= 0,
\end{align*}
\]  
\[\text{(4.102)}\]

or in matrix form as

\[
\frac{d}{d\zeta} \begin{pmatrix} A_+ \\ A_- \end{pmatrix} = M \begin{pmatrix} A_+ \\ A_- \end{pmatrix},
\]  
\[\text{(4.104)}\]

where

\[
M = \begin{pmatrix} i\delta & iD \\ -iD^* & -i\delta \end{pmatrix}.
\]  
\[\text{(4.105)}\]

In what follows \(D\) is assumed to be a real constant. For this case the solution is given by

\[
\begin{pmatrix} A_+(\zeta) \\ A_- (\zeta) \end{pmatrix} = \exp (M\zeta) \begin{pmatrix} A_+ (0) \\ A_- (0) \end{pmatrix}.
\]  
\[\text{(4.106)}\]

For a complex number \(\zeta\) and for a unit vector \(\hat{n}\) (i.e., \(\hat{n} \cdot \hat{n} = 1\)), the following holds

\[
\exp (\zeta \sigma \cdot \hat{n}) = \cosh \zeta + \sigma \cdot \hat{n} \sinh \zeta,
\]  
\[\text{(4.107)}\]

where \(\sigma = (\sigma_x, \sigma_y, \sigma_z)\) is the vector of Pauli matrices, which are given by

\[
\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.
\]  
\[\text{(4.108)}\]

Thus by expressing \(M\zeta\) as

\[
M\zeta = \gamma \zeta (\sigma_x, \sigma_y, \sigma_z) \cdot \begin{pmatrix} 0 \\ \frac{\gamma}{\delta} \frac{1}{\gamma} \\ \frac{\gamma}{\delta} \frac{i}{\gamma} \end{pmatrix},
\]  
\[\text{(4.109)}\]

where

\[
\gamma = \sqrt{D^2 - \delta^2},
\]  
\[\text{(4.110)}\]

one finds that

\[
\exp (M\zeta) = \begin{pmatrix} \cosh (\gamma \zeta) + \frac{i\delta \sinh (\gamma \zeta)}{\gamma} & iD \sinh (\gamma \zeta) \\ \frac{iD^* \sinh (\gamma \zeta)}{\gamma} & \cosh (\gamma \zeta) - \frac{i\delta \sinh (\gamma \zeta)}{\gamma} \end{pmatrix}.
\]  
\[\text{(4.111)}\]
4.4 Problems

1. An object having height $y_1$ is placed a distance $s_1$ from a lens having focal length $f$, and an image having height $y_2$ is formed at a distance $s_2$ from the lens. Find a relation between $s_1$, $s_2$ and $f$ and calculate the magnification $M = y_2/y_1$.

2. Consider an incident laser spot having a radius $R_{in}$ and a small divergence angle $\theta_{in}$. Employ paraxial ray optics to estimate the output radius $R_{out}$ and output divergence angle $\theta_{out}$ at the plane $z_{out}$ for the following cases:
   a) The spot is focused by illuminating a lens having focal length $f$ and $z_{out}$ is taken to be the location of the focal plane of the lens.
   b) The spot is collimated by locating it at the focal plane of a lens having focal length $f$ and $z_{out}$ is taken to be the location of the rear plane of the lens.
   c) The spot is expanded by employing two lenses having focal lengths $f_1$ and $f_2$ respectively placed a distance $d$ one from the other.

3. A GRIN lens having length $z$, maximum refractive index $n_0$ and pitch $p$ [see Eqs. (4.49) and (4.55)] is employed for imaging. An object is located at a distance $s_2$ from one interface of the GRIN lens, and its image is created at at distance $s_1$ from the opposite interface. Express the magnification $M$ in terms of $z$, $n_0$, $p$ and $s_1$.

4. Consider the task of coupling a collimated laser beam of wavelength $\lambda$ having characteristic mode radius $w_L$ into a single mode optical fiber having characteristic mode radius $w_F$ and refractive index $n$. The task is performed by focusing the laser beam into the fiber using a lens. What is the optimized choice for the value of the focal length of the lens $f$?

5. Consider the cavity seen in Fig. 4.2, which is made of two flat mirrors and two lenses both having focal distance $f$. The distance between the left (right) mirror and the left (right) lens is $z_1$, and the distance between the lenses is $2z_2$. All elements share the same optical axis. Under what condition the cavity is stable?

6. An optical cavity of length $d$ is formed between two concave mirrors facing each other (both mirrors are centered with respect to the optical axis of the cavity, and both are normal to the optical axis). The left (right) mirror has a radius of curvature $R_1$ ($R_2$). Find the location and spot size of the waist of a gaussian mode trapped inside the cavity.

7. Consider a gaussian beam in free space having wavelength $\lambda$. A lens having focal distance $f$ is positioned at the location of the waist of the beam, which has a spot size $w_0$. Calculate the spot size $w_L$ of the new waist that is generated by the focusing effect of the lens and the distance $d_L$ between it and the lens.
4.5 Solutions

1. The ABCD matrix is given by [see Eqs. (4.2), (4.3) and (4.10)]

\[
\begin{pmatrix}
A & B \\
C & D
\end{pmatrix} = \begin{pmatrix} 1 & s_2 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ -\frac{1}{f} & 1 \end{pmatrix} \begin{pmatrix} 1 & s_1 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} \frac{f-s_2}{f} & \frac{s_1 f - s_1 s_2 + s_2 f}{f} \\ -\frac{1}{f} & \frac{-s_1 + f}{f} \end{pmatrix}.
\]

(4.112)

Imaging occurs when \( B = 0 \) (explain why), i.e. when

\[
\frac{1}{s_1} + \frac{1}{s_2} = \frac{1}{f}.
\]

(4.113)

When \( B = 0 \) the matrix becomes

\[
\begin{pmatrix}
A & B \\
C & D
\end{pmatrix} = \begin{pmatrix} M & 0 \\ 0 & \frac{1}{M} \end{pmatrix},
\]

(4.114)

where

\[
M = -\frac{s_2}{s_1}
\]

(4.115)

is the magnification.

2. With the help of Eqs. (4.2), (4.3) and (4.10) one finds that:

a) For the case of focusing \( R_{\text{out}} \) and \( \theta_{\text{out}} \) are given by

\[
\begin{pmatrix} R_{\text{out}} \\ \theta_{\text{out}} \end{pmatrix} = \begin{pmatrix} 1 & f \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ -\frac{1}{f} & 1 \end{pmatrix} \begin{pmatrix} R_{\text{in}} \\ \theta_{\text{in}} \end{pmatrix} = \begin{pmatrix} f \theta_{\text{in}} \\ \frac{-R_{\text{in}} + f \theta_{\text{in}}}{f} \end{pmatrix},
\]

(4.116)
b) For the case of collimating $R_{\text{out}}$ and $\theta_{\text{out}}$ are given by

\[
\begin{bmatrix}
R_{\text{out}} \\
\theta_{\text{out}}
\end{bmatrix} = \begin{bmatrix} 1 & 0 \\ -\frac{1}{f} & 1 \end{bmatrix} \begin{bmatrix} 1 & f \\ 0 & 1 \end{bmatrix} \begin{bmatrix} R_{\text{in}} \\
\theta_{\text{in}}
\end{bmatrix}
= \begin{bmatrix} R_{\text{in}} + f\theta_{\text{in}} \\
-\frac{R_{\text{in}} - f\theta_{\text{in}}}{f} \end{bmatrix},
\]

(4.117)

c) For the case spot expansion the ABCD matrix is given by

\[
\begin{bmatrix} A & B \\ C & D \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ -\frac{1}{f_2} & 1 \end{bmatrix} \begin{bmatrix} 1 & d \\ 0 & 1 \end{bmatrix} \begin{bmatrix} -\frac{f_1 + d}{f_2} \\ -\frac{f_1 f_2 + f_2}{f_2} \end{bmatrix}

= \begin{bmatrix} f_1 + f_2 \\ -\frac{d - f_2}{f_1} \end{bmatrix}.
\]

(4.118)

Collimating occurs when $C = 0$, i.e. when

\[d = f_1 + f_2.\]

(4.119)

For this case $R_{\text{out}}$ and $\theta_{\text{out}}$ are given by

\[
\begin{bmatrix}
R_{\text{out}} \\
\theta_{\text{out}}
\end{bmatrix} = \begin{bmatrix} \frac{f_2 f_1}{f_2} & f_1 + f_2 \\ 0 & -\frac{f_2}{f_2} \end{bmatrix} \begin{bmatrix} R_{\text{in}} \\
\theta_{\text{in}}
\end{bmatrix} = \begin{bmatrix} -\frac{f_2 R_{\text{in}}}{f_1} + \theta_{\text{in}} d \\
\frac{1}{f_2} \theta_{\text{in}} \end{bmatrix}.
\]

(4.120)

3. The ABCD matrix is given by [see Eqs. (4.2), (4.3), (4.6) and (4.50)]

\[
\begin{bmatrix} A & B \\ C & D \end{bmatrix} = \begin{bmatrix} 1 & s_1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & n_0 \end{bmatrix}
\times \begin{bmatrix} \cos (gz) & \frac{1}{g} \sin (gz) \\ -g \sin (gz) & \cos (gz) \end{bmatrix}
\times \begin{bmatrix} 1 & s_2 \\ 0 & 1 \end{bmatrix}
= \cos (gz) \begin{bmatrix} 1 - n_0 g s_1 \tan (gz) & \frac{n_0 g (s_1 + s_2) + (1 - n_0 g^2 s_1 s_2) \tan (gz)}{g n_0} \\ -n_0 g \tan (gz) & -n_0 g s_2 \tan (gz) + 1 \end{bmatrix}.
\]

(4.121)

where [see Eq. (4.55)]

\[g = \frac{2\pi}{\phi}.\]

(4.122)

The displacement $s_2$ is determined from the imaging condition
4.5. Solutions

\[ 0 = B = \frac{n_0 g (s_1 + s_2) + (1 - n_0^2 g^2 s_1 s_2) \tan (gz)}{gn_0} , \quad (4.123) \]

which leads to

\[ s_2 = \frac{1}{gn_0} \frac{n_0 g s_1 + \tan (gz)}{n_0 g s_1 \tan (gz) - 1} . \quad (4.124) \]

For this case the ABCD matrix becomes [see Eq. (4.114)]

\[
\begin{pmatrix}
A & B \\
C & D
\end{pmatrix} = \begin{pmatrix}
M & 0 \\
-\frac{1}{f} & \frac{1}{M}
\end{pmatrix} ,
\]

(4.125)

where the focal distance \( f \) is given by

\[ f = \frac{1}{n_0 g \sin (gz)} , \quad (4.126) \]

and where the magnification \( M \) is given by

\[ M = \cos \frac{2\pi z}{p} - \frac{2\pi n_0 s_1}{p} \sin \frac{2\pi z}{p} . \quad (4.127) \]

4. Coupling into the fiber is optimized by choosing a lens which focuses the laser beam into a spot having characteristic mode radius as close as possible to \( w_F \). The ABCD matrix associated with the inverse transformation from the fiber edge, which is taken to be the input plane, to the plane beyond the lens, which is taken to be the output plane, is given by

\[
\begin{pmatrix}
A & B \\
C & D
\end{pmatrix} = \begin{pmatrix}
1 & 0 \\
-\frac{1}{f} & 1
\end{pmatrix} \begin{pmatrix}
1 & f \\
0 & 1
\end{pmatrix} \begin{pmatrix}
1 & 0 \\
0 & n
\end{pmatrix}
\]

\[ = \begin{pmatrix}
1 & fn \\
-\frac{1}{f} & 0
\end{pmatrix} . \quad (4.128) \]

Note that it is assumed that the lens is positioned at a distance \( f \) from the fiber end, in order to obtain collimation. The input complex beam parameter \( q_{in} \) is given by [see Eq. (4.65)]

\[ q_{in} = -i \frac{\pi w_F^2 n}{\lambda} , \quad (4.129) \]

and the output complex beam parameter \( q_{out} \) is given by

\[ q_{out} = -i \frac{\pi w_F^2}{\lambda} . \quad (4.130) \]

Using Eq. (4.66) one finds that
\[ q_{\text{out}} = \frac{Aq_i + B}{Cq_i + D}, \quad (4.131) \]

thus
\[ i \frac{\pi w_i^2}{\lambda} = f \left( i \frac{f}{\pi w_F^2} + 1 \right). \quad (4.132) \]

The assumption that \( w_F \ll \sqrt{f\lambda} \) implies that the real part of the above equation is much smaller than the imaginary part. By neglecting the real part one finds that
\[ f = \frac{\pi w_L w_F}{\lambda}, \quad (4.133) \]

or in terms of the diameters \( 2w_L \) and \( 2w_F \)
\[ f = \frac{\pi (2w_L)(2w_F)}{4\lambda}. \quad (4.134) \]

5. The ABCD matrix corresponding to a single trip from the left mirror to the right one is given by [see Eqs. (4.3) and (4.10)]
\[ M = \begin{pmatrix} 1 & z_1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ -\frac{1}{f} & 1 \end{pmatrix} \begin{pmatrix} 1 & 2z_2 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ -\frac{1}{f} & 1 \end{pmatrix} \begin{pmatrix} 1 & z_1 \\ 0 & 1 \end{pmatrix}. \quad (4.135) \]

The stability condition reads [see Eq. (4.25) and note that \( \det M = 1 \)]
\[ \left| \frac{\text{Tr}(M)}{2} \right| = \left| 1 + \frac{2z_1 z_2}{f} \left( \frac{1}{f} - \frac{z_1 + z_2}{z_1 z_2} \right) \right| \leq 1, \quad (4.136) \]

and thus
\[ f \geq \frac{z_1 z_2}{z_1 + z_2}, \quad (4.137) \]

or
\[ \frac{1}{f} \leq \frac{1}{z_1} + \frac{1}{z_2}. \quad (4.138) \]

6. Let \( z_1 \quad (z_2 = z_1 + d) \) be the location of the left (right) mirror along the optical axis, and let \( z = 0 \) be the location of the waist. A gaussian mode trapped inside the cavity has a fixed position-dependent radius of curvature \( R(z) \). In general, in the paraxial approximation back reflection by a concave mirror having radius of curvature \( R \) is described by an ABCD matrix given by (recall that the focal length \( f \) of a spherical mirror having radius \( R \) is \( f = R/2 \))
\[ \begin{pmatrix} A & B \\ C & D \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ \frac{-2}{R} & 1 \end{pmatrix}, \quad (4.139) \]
and thus the complex beam parameter $q$ of a back-reflected gaussian beam $q_{\text{out}}$ is related to the parameter $q$ of an incident beam $q_{\text{in}}$ by [see Eq. (4.66)]

\[
\frac{1}{q_{\text{out}}} = \frac{1}{q_{\text{in}}} - \frac{2}{R}.
\]  

(4.140)

Recall that real \( \left( \frac{1}{q} \right) = \frac{1}{R(z)} \) [see Eq. (4.65)], and thus for a cavity mode having a fixed position-dependent radius of curvature $R(z)$ the following must hold

\[
R(z_1) = -R_1,
\]

(4.141)

\[
R(z_2) = R_2,
\]

(4.142)

where [see Eq. (4.78)]

\[
R(z) = z \left( 1 + \frac{|q_0|^2}{z^2} \right),
\]

(4.143)

and

\[
q_0 = q(z = 0) = \frac{n_0 k_0 w_0^2}{2},
\]

(4.144)

where $w_0$ is the spot size at the the location of the waist, i.e. at $z = 0$. Equations (4.141) and (4.142) together with the requirement that $z_2 - z_1 = d$ can be expressed as

\[
d - \frac{|q_0|^2}{2} \cdot \frac{d}{z_0} = 2R_0,
\]

(4.145)

\[-z_0 - \frac{|q_0|^2 z_0}{2} \cdot \frac{d}{z_0} = R_m,
\]

(4.146)

where

\[
R_0 = \frac{R_1 + R_2}{2},
\]

(4.147)

\[
R_m = \frac{R_1 - R_2}{2},
\]

(4.148)

\[
z_0 = \frac{z_1 + z_2}{2},
\]

(4.149)

and thus

\[
z_0 = \frac{d}{2} \frac{R_m}{R_0 - d},
\]

(4.150)

and

\[
|q_0|^2 = \left( \frac{d}{2} \right)^2 \left( 1 + \frac{R_m}{z_0} \right) \left( 1 - \frac{R_m^2}{(R_0 - d)^2} \right),
\]

(4.151)
or in terms of \( R_1 \) and \( R_2 \)
\[
|q_0|^2 = \frac{d(R_1 - d)(R_2 - d)(R_1 + R_2 - d)}{(R_1 + R_2 - 2d)^2},
\]
(4.152)
\[
z_1 = -\frac{d(R_2 - d)}{R_1 + R_2 - 2d},
\]
(4.153)
\[
z_2 = \frac{d(R_1 - d)}{R_1 + R_2 - 2d},
\]
(4.154)
and thus [see Eq. (4.144)]
\[
w_0 = \left[ \frac{4d(R_1 - d)(R_2 - d)(R_1 + R_2 - d)}{n_0^2 k_0^2 (R_1 + R_2 - 2d)^2} \right]^{1/4}.
\]
(4.155)

7. The ABCD matrix corresponding to the transformation from the location of the lens to the new waist is given by [see Eqs. (4.3) and (4.10)]
\[
\begin{pmatrix}
A & B \\
C & D
\end{pmatrix} = \begin{pmatrix} 1 & d_f \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ -\frac{1}{f} & 1 \end{pmatrix} = \begin{pmatrix} 1 & -\frac{d_f}{f} \\ -\frac{1}{f} & 1 \end{pmatrix}.
\]
(4.156)
The following holds [see Eq. (4.66)]
\[
q(d_f) = \frac{Aq_0 + B}{Cq_0 + D},
\]
(4.157)
where [see Eqs. (4.65), (4.78) and (4.79) and recall that \( k_0 = 2\pi/\lambda \)]
\[
\frac{1}{q_0} = \frac{i\lambda}{\pi w_0^2},
\]
(4.158)
\[
\frac{1}{q(d_f)} = \frac{i\lambda}{\pi w_f^2},
\]
(4.159)
and thus Eq. (4.157) becomes
\[
\frac{1}{2ic} = \frac{-\frac{1}{f} \left( 1 - \frac{d_f}{f} \right) + df \left( \frac{\lambda}{\pi w_f^2} \right)^2 - \frac{i\lambda}{\pi w_0^2}}{\frac{1}{f^2} + \left( \frac{\lambda}{\pi w_f^2} \right)^2}.
\]
(4.160)
The real part of Eq. (4.160) yields
\[
d_f = \frac{f}{1 + \eta^2},
\]
(4.161)
and the imaginary part yields
\[
w_f^2 = \frac{w_0^2}{1 + \eta^{-2}},
\]
(4.162)
where
\[
\eta = \frac{f\lambda}{\pi w_0^2}.
\]
(4.163)
5. Scalar Diffraction Theory

Consider the case where sources located in the left half space \( z < 0 \) generate a monochromatic electromagnetic field at angular frequency \( \omega \). The right half space \( z > 0 \) is assumed to be a vacuum free of any matter and sources. The theory of scalar diffraction allows evaluating the electromagnetic field in the right half space \( z > 0 \) in terms of the field in the plane \( z = 0 \).

5.1 Angular Spectrum

In the right half space \( z > 0 \) all components of the electric and magnetic fields satisfy the Helmholtz equation (2.153), which is given by

\[
(\nabla^2 + k^2) u = 0 ,
\]  

(5.1)

where

\[
k = \frac{\omega}{c} .
\]  

(5.2)

For any value of \( z \) the function \( u(x, y, z) \) can be Fourier expanded in the lateral \( xy \) plane. The two-dimensional Fourier transformed function \( u(k_x, k_y, z) \) is given by

\[
u (k_x, k_y, z) = \mathcal{F}(u(x, y, z)) ,
\]  

(5.3)

where

\[
\mathcal{F}(u(x, y, z)) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \text{d}x \text{d}y \ u(x, y, z) e^{-i(k_x x + k_y y)} .
\]  

(5.4)

Claim. The inverse Fourier transform is given by

\[
\mathcal{F}^{-1}(u(k_x, k_y, z)) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \text{d}k_x \text{d}k_y \ u(k_x, k_y, z) e^{i(k_x x + k_y y)} .
\]  

(5.5)
Proof. With the help of the identity
\[ \int_{-\infty}^{\infty} \, d\kappa \, e^{i\kappa s} = 2\pi \delta(s) \, , \]  
(5.6)
one obtains [see Eqs. (5.4) and (5.5)]
\[ \mathcal{F}^{-1} \left( \mathcal{F} \left( u(x', y', z) \right) \right) \]
\[ = \frac{1}{4\pi^2} \int_{-\infty}^{\infty} \, dx \int_{-\infty}^{\infty} \, dy \, u(x, y, z) \int_{-\infty}^{\infty} \, dk_x \, e^{ik_x(x'-x)} \int_{-\infty}^{\infty} \, dk_y \, e^{ik_y(y'-y)} \]
\[ = u(x', y', z) \, . \]  
(5.7)

The two-dimensional Fourier transformed function in the plane \( z = 0 \) is denoted by
\[ U(k_x, k_y) = u(k_x, k_y, z = 0) \, , \]  
(5.8)
where [see Eq. (5.4)]
\[ U(k_x, k_y) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \, dx \, dy \, u(x, y, 0) \, e^{-i(k_x x + k_y y)} \]  
(5.9)

Claim. The following holds
\[ u(k_x, k_y, z) = U(k_x, k_y) \, e^{ik_z z} \, , \]  
(5.10)
where
\[ k_z = \sqrt{k^2 - k_x^2 - k_y^2} \, . \]  
(5.11)

Proof. By substituting \( u(x, y, z) = \mathcal{F}^{-1} \left( u(k_x, k_y, z) \right) \) [see Eq. (5.5)] into the Helmholtz equation (5.1) one obtains
\[ \frac{1}{2\pi} \int_{-\infty}^{\infty} \, dk_x \, dk_y \, e^{i(k_x x + k_y y)} \left( k_z^2 + \frac{\partial^2}{\partial z^2} \right) u(k_x, k_y, z) = 0 \, , \]  
(5.12)
and thus
\[ \left( k_z^2 + \frac{\partial^2}{\partial z^2} \right) u(k_x, k_y, z) = 0 \, . \]  
(5.13)
The solution of (5.13) leads to Eq. (5.10).
Note that when \( k_x^2 + k_y^2 < k^2 \) the term \( e^{ik_z z} \) represents a plane wave, whereas when \( k_x^2 + k_y^2 > k^2 \) it represents an evanescent wave. Thus Eq. (5.10) implies that Fourier components in the plane \( z = 0 \) having high spatial frequency, for which \( k_x^2 + k_y^2 > k^2 \), do not reach the far field (i.e. values of \( z \) much larger than a single wavelength) since they exponentially decay as a function of \( z \).

The expression given by Eq. (5.3) allows expressing \( U(k_x, k_y) \) in terms of \( u \) in the plane \( z = 0 \). As is shown below, \( U(k_x, k_y) \) can alternatively be expressed in terms of the normal derivative \( \partial u / \partial z \) in the plane \( z = 0 \).

Claim. The following holds

\[
U(k_x, k_y) = \frac{1}{2\pi i k_z} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} dx dy \frac{\partial u}{\partial z} \bigg|_{z=0} e^{-i(k_x x + k_y y)} .
\] (5.14)

Proof. The following holds [see Eqs. (5.5) and (5.10)]

\[
u(x, y, z) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} dk_x dk_y U(k_x, k_y) e^{ik \cdot r},
\] (5.15)

where \( r = (x, y, z) \), \( k = (k_x, k_y, k_z) \), and \( k_z \) is given by Eq. (5.11). Taking the derivative with respect to \( z \) and setting \( z = 0 \) lead to

\[
\frac{\partial u}{\partial z} \bigg|_{z=0} = \frac{1}{2\pi} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} dk_x dk_y i k_z U(k_x, k_y) e^{i(k_x x + k_y y)}. 
\] (5.16)

The expression given by Eq. (5.14) is obtained by applying the two-dimensional Fourier transform (5.4), i.e. by multiplying by \( e^{-i(k'_x x + k'_y y)} \), integrating over \( x \) and \( y \) and employing the identity (5.6).

The above results can be employed in order to express \( u \) in terms of either its value in the plane \( z = 0 \) or in terms of its normal derivative in the plane \( z = 0 \) [see Eqs. (5.18) and (5.19) below, respectively]

Claim. The following holds

\[
u(r') = u_1(r') = u_2(r') ,
\] (5.17)

where

\[
u_1(r') = 2 \int_{-\infty}^{\infty} dx'' \int_{-\infty}^{\infty} dy'' u(r'') \frac{\partial g(r' - r'')}{\partial z''},
\] (5.18)
\[ u_2 (r') = -2 \int_{-\infty}^{\infty} dx'' \int_{-\infty}^{\infty} dy'' \frac{\partial u (r'')}{\partial z''} g (r' - r'') , \quad (5.19) \]

where \( r' = (x', y', z') \), \( r'' = (x'', y'', z'') \), in both Eqs. (5.18) and (5.19) the integrals are evaluated in the plane \( z'' = 0 \), the function \( g \) is given by

\[ g (r) = \frac{i}{8 \pi^2} \int_{-\infty}^{\infty} dk_x \int_{-\infty}^{\infty} dk_y \frac{e^{ik \cdot r}}{k_z} , \quad (5.20) \]

and \( k_z \) is given by Eq. (5.11).

**Proof.** With the help of Eqs. (5.9) and (5.15) one finds that

\[ u (r') = \frac{1}{4 \pi^2} \int_{-\infty}^{\infty} dx'' \int_{-\infty}^{\infty} dy'' \int_{-\infty}^{\infty} dk_x \int_{-\infty}^{\infty} dk_y \left. u (x'', y'', 0) e^{ik \cdot r'} \right|_{z'' = 0} \tag{5.21} \]

Similarly, Eqs. (5.14) and (5.15) yield

\[ u (r') = \frac{1}{4 \pi^2} \int_{-\infty}^{\infty} dx'' \int_{-\infty}^{\infty} dy'' \frac{\partial u (r'')}{\partial z''} \int_{-\infty}^{\infty} dk_x \int_{-\infty}^{\infty} dk_y \frac{e^{ik \cdot r' - i(k_z x'' + k_y y'')}}{ik_z} \left. \right|_{z'' = 0} \tag{5.22} \]

The expression for \( u_1 (r') \) (5.18) and \( u_2 (r') \) (5.19) can be further simplified with the help of the so-called Weyl’s plane waves expansion of a scalar spherical wave [see Eq. (5.34) below]

**Claim.** For \( z \geq 0 \) the function \( g (r) \) [see Eq. (5.20)] can be expressed as

\[ g (r) = u_S (r) , \quad (5.23) \]

where the scalar spherical wave \( u_S (r) \) is defined by

\[ u_S (r) = -\frac{e^{ikr}}{4 \pi r} , \quad (5.24) \]

\( k \) is given by Eq. (5.2) and \( r = \sqrt{x^2 + y^2 + z^2} \).
Proof. The two-dimensional Fourier transform of $u_S(r)$ in the plane $z = 0$, which is denoted by $U_S(k_x, k_y)$, is given by [see Eq. (5.9)]

$$U_S(k_x, k_y) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} dx dy \ u_S(x, y, 0) e^{-(k_x x + k_y y)}$$

$$= -\frac{1}{8\pi^2} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} dx dy \ \frac{e^{ik\sqrt{x^2+y^2}}}{\sqrt{x^2+y^2}} e^{-(k_x x + k_y y)} .$$

(5.25)

The coordinate transformation

$$x = \rho \cos \theta ,$$

(5.26)

$$y = \rho \sin \theta ,$$

(5.27)

$$k_x = \kappa \cos \varphi ,$$

(5.28)

$$k_y = \kappa \sin \varphi ,$$

(5.29)

together with the identity

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} d\theta e^{-iA \cos \theta} = J_0(A) ,$$

(5.30)

where $J_0$ is the Bessel function of the first kind, yield

$$U_S(k_x, k_y) = -\frac{1}{4\pi^2} \int_{0}^{\infty} d\rho \ e^{ik\rho} J_0(\kappa \rho) \ .$$

(5.31)

The identity

$$\int_{0}^{\infty} d\rho \ e^{ik\rho} J_0(\kappa \rho) = \frac{1}{i\sqrt{k^2 - \kappa^2}} ,$$

(5.32)

leads to

$$U_S(k_x, k_y) = -\frac{1}{4\pi i k_z} .$$

(5.33)

where $k_z$ is given by Eq. (5.11). The above result together with Eqs. (5.5) and (5.10) lead to the Weyl's expansion.
\[
\frac{e^{ikr}}{r} = \frac{1}{2\pi i} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} dk_x dk_y \frac{e^{ikr}}{k_z},
\]

(5.34)

and hence Eq. (5.23) holds [see Eq. (5.20)].

5.2 Rayleigh-Sommerfeld, Kirchhoff, Fresnel and Fraunhofer Diffraction Integrals

Combining Eqs. (5.18) and (5.23) leads to the so-called Rayleigh-Sommerfeld first diffraction integral

\[
u_1 (r') = -\frac{1}{2\pi} \int_{-\infty}^{\infty} dx'' \int_{-\infty}^{\infty} dy'' u(r'') \frac{\partial}{\partial z''} \left( \frac{e^{ikr}}{r} \right),
\]

(5.35)

whereas the so-called second Rayleigh-Sommerfeld diffraction integral is obtained by combining Eqs. (5.19) and (5.23)

\[
u_2 (r') = \frac{1}{2\pi} \int_{-\infty}^{\infty} dx'' \int_{-\infty}^{\infty} dy'' \frac{\partial u(r'')}{\partial z''} \frac{e^{ikr}}{r},
\]

(5.36)

where

\[r = |r' - r''|.
\]

(5.37)

The Kirchhoff diffraction integral \(u_K (r')\) is defined by [see Eqs. (5.18) and (5.19) and compare with Eq. (5.85) below]

\[
u_K (r') = \frac{\nu_1 (r') + \nu_2 (r')}{2} = \int_{-\infty}^{\infty} dx'' \int_{-\infty}^{\infty} dy'' \left( u(r'') \frac{\partial g(r' - r'')}{\partial z''} - \frac{\partial u(r'')}{\partial z''} g(r' - r'') \right).
\]

(5.38)

5.2.1 The Limit of Geometrical Optics

Consider in general an integral \(I\), which is given by

\[
I = \int_{-\infty}^{\infty} dx \ h(x) e^{ikf(x)},
\]

(5.39)

where the functions \(h(x)\) and \(f(x)\) and the coefficient \(k\) are all real. Let \(I_a (x_0, \delta)\) be the contribution to \(I\) from the interval \([x_0 - \delta, x_0 + \delta]\), i.e.
5.2. Rayleigh-Sommerfeld, Kirchhoff, Fresnel and Fraunhofer Diffraction Integrals

\[ I_d (x_0, \delta) = \int_{x_0 - \delta}^{x_0 + \delta} dx \ h(x) e^{ikf(x)}. \]  \hspace{1cm} (5.40)

For sufficiently small \( \delta \) the following approximations are expected to hold inside the interval \([x_0 - \delta, x_0 + \delta]\)

\[ h(x) \simeq h(x_0), \]  \hspace{1cm} (5.41)
\[ f(x) \simeq f(x_0) + (x - x_0) f'(x_0), \]  \hspace{1cm} (5.42)

and thus

\[ I_d (x_0, \delta) \simeq 2h(x_0) e^{ikf(x_0)} \sin \left( \frac{\delta k f'(x_0)}{k f''(x_0)} \right). \]  \hspace{1cm} (5.43)

The above result (5.43) indicates that in the limit of geometrical optics, i.e. when \( k \) is large, the main contribution to the integral \( I \) comes from regions near points at which the phase factor \( k f(x) \) in Eq. (5.39) is locally stationary, i.e. points \( x_s \) such that \( f'(x_s) = 0. \)

**Exercise 5.2.1.** Calculate \( I \) in the limit \( k \to \infty \) for the case where \( f(x) \) has a single stationary point.

**Solution 5.2.1.** By employing the Taylor expansion near the stationary point \( x_s \)

\[ f(x) = f(x_s) + \frac{(x - x_s)^2}{2} f''(x_s) + \cdots, \]  \hspace{1cm} (5.44)

where prime denotes a derivative with respect to \( x \), and the variable transformation \( s = x - x_s \), one obtains in the limit \( k \to \infty \)

\[ I \simeq e^{ikf(x_s)} h(x_s) \int_{-\infty}^{\infty} ds \ e^{is^2 / 2 f''(x_s)}. \]  \hspace{1cm} (5.45)

With the help of the identity

\[ \int_{-\infty}^{\infty} dx \ \exp \left( -ax^2 + bx + c \right) = \sqrt{\frac{\pi}{a}} e^{\frac{1}{4} \frac{b^2 - 4ac}{a}}, \]  \hspace{1cm} (5.46)

this becomes

\[ I \simeq h(x_s) \sqrt{\frac{2\pi i}{k f''(x_s)}} e^{ikf(x_s)}. \]  \hspace{1cm} (5.47)

The above result (5.47) is known as the stationary phase approximation. The characteristic width \( \Delta x \) of the interval around the stationary point \( x_s \) that is responsible for the dominant contribution to the value of the integral \( I \) is given by

\[ \Delta x = \frac{1}{\sqrt{k f''(x_s)}}. \]  \hspace{1cm} (5.48)
Exercise 5.2.2. Calculate the Rayleigh-Sommerfeld first diffraction integral (5.35) in the limit \( k \to \infty \).

Solution 5.2.2. With the help of Eqs. (5.35) and (5.53) one obtains

\[
 u_1(r') = \frac{1}{2\pi} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \frac{\partial}{\partial z''} u(r'') \left( \frac{e^{ikr} (ikr - 1) z'}{r^3} \right),
\]

where

\[
 r = z' \sqrt{1 + \frac{(x' - x'')^2 + (y' - y'')^2}{z'^2}},
\]

thus in the limit \( k \to \infty \) [see Eq. (5.47)]

\[
 u_1(r') = u(x', y', 0) e^{ikz'}. \tag{5.51}
\]

The above represents the limit of geometrical optics.

5.3 The Fresnel and Fraunhofer Diffraction Integrals

The Rayleigh-Sommerfeld and Kirchhoff diffraction integrals [see Eqs. (5.35), (5.36) and (5.38)] can be further simplified by applying approximations. The factor multiplying \( u(r'') \) in Rayleigh-Sommerfeld first diffraction integral (5.35) is given by

\[
 \frac{\partial}{\partial z''} \left( \frac{e^{ikr}}{r} \right) = \frac{e^{ikr} (ikr - 1)}{r^2} \frac{\partial r}{\partial z''}, \tag{5.52}
\]

thus for \( z'' = 0 \) [see Eq. (5.37)]

\[
 \frac{\partial}{\partial z''} \left( \frac{e^{ikr}}{r} \right) = -\frac{e^{ikr} (ikr - 1) z'}{r^3}, \tag{5.53}
\]

where

\[
 r = z' \sqrt{1 + \frac{(x' - x'')^2 + (y' - y'')^2}{z'^2}}. \tag{5.54}
\]

The so-called Fresnel diffraction integral [see Eq. (5.56) below] is obtained by assuming (a) the far field limit, i.e. \( kr \gg 1 \), and the paraxial case, for which it is assumed that \( r \approx z' \), i.e.

\[
 \frac{(x' - x'')^2 + (y' - y'')^2}{z'^2} \ll 1. \tag{5.55}
\]
These assumptions together with Eqs. (5.35) and (5.53) lead to the Fresnel diffraction integral

\[ u_1(r') = \frac{ike^{ikz'}}{2\pi z'} \int_{-\infty}^{\infty} dx'' \int_{-\infty}^{\infty} dy'' u(r'') e^{i \frac{(x'-x'')^2 + (y'-y'')^2}{2z'}}. \] (5.56)

**Exercise 5.3.1.** Consider the case where the scalar \( u(r'') \) in the plane \( z'' = 0^+ \) is taken to be given by

\[ u(x'',y'',z''=0^+) = u_{\text{inc}}(x'',y'') t(x'',y''), \] (5.57)

where for normal incident \( u_{\text{inc}}(x'',y'') \) is assumed to be a positive constant denoted by \( u_0 \) and the aperture transmission \( t(x'',y'') \) is given by

\[ t(x'',y'') = \begin{cases} 1 & x'' < 0 \\ 0 & x'' \geq 0 \end{cases}. \] (5.58)

Employ the Fresnel diffraction integral (5.56) in order to roughly estimate what region in the half space \( z' > 0 \) remains dark (i.e. the region where \( |u_1(r')| \ll u_0 \)).

**Solution 5.3.1.** In general, the Fresnel diffraction integral (5.56) together with Eq. (5.48) imply that the value of \( u_1(r') \) is determined by a region in the plane \( z'' = 0 \) centered at \((x', y') = (x'', y'')\) and having characteristic widths in the \( x \) and \( y \) directions roughly given by \( \Delta x'' = \Delta y'' = \sqrt{2z''/k}. \)

For the current case under consideration \( u(x'',y'',z''=0^+) \) vanishes when \( x'' \geq 0 \), and consequently it is expected that \( |u_1(r')| \ll u_0 \) provided that \( x'' \gtrsim \Delta x'' = \sqrt{z''/k}. \)

Fraunhofer diffraction integral [see Eq. (5.60) below] is derived from the additional assumption that

\[ \frac{k \left( x''^2 + y''^2 \right)}{2z'} \ll 1. \] (5.59)

In this limit Eq. (5.56) leads to the Fraunhofer diffraction integral

\[ u_1(r') = \frac{ik\Phi}{2\pi z'} \int_{-\infty}^{\infty} dx'' \int_{-\infty}^{\infty} dy'' u(r'') e^{-i \left( \kappa_x x'' + \kappa_y y'' \right)}, \] (5.60)

where

\[ \Phi = e^{ik \left( z' + \frac{x'^2 + y'^2}{2z'} \right)}, \] (5.61)

and where

\[ \kappa_x = \frac{kx'}{z'}, \kappa_y = \frac{ky'}{z'}. \] (5.62)

As can be seen from the comparison between Eqs. (5.4) and (5.60), the Fraunhofer diffraction integral is proportional to the two-dimensional Fourier transform of \( u(r'') \).
5.4 Imaging

Consider a lens having transmission \( t_L(x, y) \), which is taken to be given by

\[
\begin{align*}
t_L(x, y) &= e^{-\frac{ik(x^2 + y^2)}{2f}} P_L(x, y),
\end{align*}
\]

where \( f \) is the focal length of the lens. The so-called pupil function \( P_L(x, y) \), which for the present case is taken to be given by

\[
\begin{align*}
P_L(x, y) &= \begin{cases} 1 & |x| \leq A_L \text{ and } |y| \leq A_L, \\
0 & \text{otherwise},
\end{cases}
\end{align*}
\]

accounts for the finite size of the lens. The lens is placed in the plane \( z = 0 \).

An aperture having transmission \( t_A(x, y) \), which is taken to be given by

\[
\begin{align*}
t_A(x'', y'') &= \delta(x'' - x_0) \delta(y'' - y_0),
\end{align*}
\]

is placed in the object plane \( z = -s_1 \). The image plane is taken to be \( z = s_2 \) on the other side of the lens. The aperture is illuminated by a plane wave having a constant amplitude in the plane \( z = -s_1 \), which is denoted by \( u_0 \).

**Exercise 5.4.1.** Calculate the scalar \( u_1(x''', y''', s_2) \) in the object plane using the Fresnel diffraction integral.

**Solution 5.4.1.** Using the Fresnel diffraction integral (5.56) one obtains for a general pupil function \( P_L(x, y) \) and a general aperture transmission \( t_A(x'', y'') \)

\[
\begin{align*}
u_1(x''', y''', s_2) &= -\frac{k^2 e^{ik(s_1+s_2)}u_0}{4\pi^2 s_1 s_2} \\
&\quad \times \int_{-\infty}^{\infty} dx' \int_{-\infty}^{\infty} dy' \int_{-\infty}^{\infty} dx'' \int_{-\infty}^{\infty} dy'' P_L(x', y') t_A(x'', y'') e^{ik\Phi},
\end{align*}
\]

where

\[
\Phi = \frac{(x' - x'')^2 + (y' - y'')^2}{s_1} + \frac{(x''' - x'')^2 + (y''' - y'')^2}{s_2} - \frac{x'^2 + y'^2}{f}
\]

\[
= \left(\frac{x'^2 + y'^2}{s_1} + \frac{x''^2 + y''^2}{s_2}\right) \left(\frac{1}{s_1} + \frac{1}{s_2} - \frac{1}{f}\right)
\]

\[
+ \frac{x'^2}{s_1} + \frac{x''^2}{s_2} - 2x' \left(\frac{x''}{s_1} + \frac{x'''}{s_2}\right) - 2y' \left(\frac{y''}{s_1} + \frac{y'''}{s_2}\right).
\]

(5.67)

When \( P_L(x, y) \) is given by Eq. (5.64) and \( t_A(x'', y'') \) by Eq. (5.65) this becomes
\[ u_1(x'', y'', s_2) = -\frac{k^2 e^{ik(s_1s_2 + s_2)\frac{u_0}{2\pi s_1 s_2}}}{A_L}\int_{-A_L}^{A_L} dx' \int_{-A_L}^{A_L} dy' e^{ik\frac{\Phi}{s_2}}, \quad (5.68) \]

where
\[ \Phi = \left( \frac{x'}{s_1} + \frac{1}{\frac{1}{s_2} - \frac{1}{f}} \right) + \frac{x''^2 + y''^2}{s_1} + \frac{x'''^2 + y'''^2}{s_2} - 2x' \left( x''' - M x_0 \right) s_2 - 2y' \left( y''' - M y_0 \right) s_2, \quad (5.69) \]

and where \( M \), which is given by
\[ M = -\frac{s_2}{s_1}, \quad (5.70) \]

is the magnification [compare with Eq. (4.115)].

When the imaging condition, which is given by [compare with Eq. (4.113)]
\[ \frac{1}{s_1} + \frac{1}{s_2} = \frac{1}{f}, \quad (5.71) \]

is satisfied the scalar \( u_1(x'', y'', s_2) \) in the object plane is found to be given by
\[
\begin{align*}
& u_1(x'', y'', s_2) \\
& = -\frac{k^2 e^{ik(s_1s_2 + s_2)}}{4\pi^2 s_1 s_2} e^{ik\left( \frac{x''^2 + y''^2}{2s_1} + \frac{x'''^2 + y'''^2}{2s_2} \right)} u_0 \\
& \times \int_{-A_L}^{A_L} dx' e^{-i\frac{k(x''' - M x_0)s_2}{s_2}} \int_{-A_L}^{A_L} dy' e^{-i\frac{k(y''' - M y_0)s_2}{s_2}} \\
& = -\frac{k^2 A_L^2 e^{ik(s_1s_2 + s_2)}}{\pi^2 s_1 s_2} e^{ik\left( \frac{x''^2 + y''^2}{2s_1} + \frac{x'''^2 + y'''^2}{2s_2} \right)} u_0 \\
& \times \text{sinc} \left( \frac{k(x''' - M x_0) A_L}{s_2} \right) \text{sinc} \left( \frac{k(y''' - M y_0) A_L}{s_2} \right),
\end{align*}
\]

where
\[ \text{sinc} q = \frac{\sin q}{q}, \quad (5.73) \]

As can be seen from Eq. (5.6), the following holds...
Chapter 5. Scalar Diffraction Theory

\[ 2\pi \delta(K) = \lim_{A \to \infty} \int_{-A}^{A} dX' e^{iKX'} = 2 \lim_{A \to \infty} A \text{sinc}(KA), \quad (5.74) \]

thus in the limit where

\[ \frac{kx''}{s_2} \gg 1, \quad (5.75) \]
\[ \frac{ky''}{s_2} \gg 1, \quad (5.76) \]

\[ u_1(x'', y'', s_2) \text{ becomes} \]

\[ u_1(x'', y'', s_2) = -k^2 e^{ik(s_1+s_2)} e^{ik\left(\frac{s_1^2+s_2^2 + x''^2+y''^2}{2s_2^2}\right)} u_0 \]
\[ \times \delta \left(\frac{k(x''-Mx_0)A_L}{s_2}\right) \delta \left(\frac{k(y''-My_0)A_L}{s_2}\right), \quad (5.77) \]

or

\[ u_1(x'', y'', s_2) = Me^{ik(s_1+s_2)} e^{ik\left(\frac{s_1^2+s_2^2 + x''^2+y''^2}{2s_2^2}\right)} u_0 \]
\[ \times \delta (x'' - Mx_0) \delta (y'' - My_0). \quad (5.78) \]

**Exercise 5.4.2.** Calculate \( u_1(x'', y'', s_2) \) for the case of a general aperture transmission \( t_A(x', y') \) that is attached directly to the lens, i.e. \( s_1 \to 0 \), and for the case where the image is generated in the focal plane of the lens, i.e. \( s_2 = f \).

**Solution 5.4.2.** With the help of Eq. (5.66) one obtains

\[ u_1(x'', y'', s_2) = -k^2 e^{ik(s_1+s_2)} e^{ik\frac{x''^2+y''^2}{2s_2^2}} u_0 \int_{-A_L}^{A_L} dx' \int_{-A_L}^{A_L} dy' e^{-ik\frac{x''^2+y''^2}{f}} \]
\[ \times \lim_{s_1 \to 0} \frac{1}{s_1} \int_{-\infty}^{\infty} dx'' \int_{-\infty}^{\infty} dy'' t_A(x'', y'') e^{ik\left(\frac{(x'-x'')^2+(y'-y'')^2}{s_1^2}\right)}, \quad (5.79) \]

thus [see Eq. (5.47)]

\[ u_1(x'', y'', s_2) = -\frac{ik e^{ik(s_1+s_2)} e^{ik\frac{x''^2+y''^2}{2s_2^2}} u_0}{s_2} t_A \left(\frac{kx''}{f}, \frac{ky''}{f}; A_L\right). \quad (5.80) \]
where

\[
t_A (k_x, k_y; A_L) = \frac{1}{2\pi} \int_{-A_L}^{A_L} dx' \int_{-A_L}^{A_L} dy' t_A (x', y') e^{-i(k_xx' + k_yy')} .
\]  
(5.81)

Note that in the limit where \( A_L \to \infty \) the term \( t_A (k_x, k_y; A_L) \) (5.81) becomes the two-dimensional Fourier transform of the aperture transmission \( t_A (x', y') \) [see Eq. (5.4)].

5.5 Problems

1. **Green’s theorem** - Show that

\[
\int_V (u \nabla^2 g - g \nabla^2 u) \, dv = \int_S \left( u \frac{\partial g}{\partial n} - g \frac{\partial u}{\partial n} \right) \, ds ,
\]

(5.82)

where \( S \) is the boundary of the volume \( V \), both \( u (r) \) and \( g (r) \) are smooth functions from \( \mathbb{R}^3 \) to \( \mathbb{C} \) and \( \partial / \partial n \) denotes a partial derivative in the outward normal direction on the boundary \( S \), i.e.

\[
\frac{\partial \psi}{\partial n} = \hat{n} \cdot \nabla u ,
\]

(5.83)

where \( \hat{n} \) is a unit vector normal to the boundary \( S \).

2. Show that

\[
(\nabla^2 + k^2) g (r) = \delta (r) ,
\]

(5.84)

where \( g (r) \) is given by Eq. (5.20).

3. **Kirchhoff diffraction integral** - Let \( u (r) \) be a solution of the Helmholtz equation (5.1) in a volume \( V \), which is bounded by the surface \( S \). Show that

\[
u (r) = \frac{1}{4\pi} \int_S \left( \frac{\partial u}{\partial n} \frac{e^{ikr}}{r} - u \frac{\partial}{\partial n} \left( \frac{e^{ikr}}{r} \right) \right) \, ds ,
\]

(5.85)

where \( \partial / \partial n \) denotes a partial derivative in the outward normal direction, \( r = |r - r'| \) and \( r' \) denotes points on the surface \( S \).

4. Consider a circular aperture of radius \( a \) normally illuminated by an incident monochromatic plane wave. Calculate the Fresnel (5.56) and Fraunhofer (5.60) diffraction integrals.

5. Calculate the Fresnel (5.56) and Fraunhofer (5.60) diffraction integrals for the case of normal incident and a rectangular aperture of sides \( 2a \) and \( 2b \) in the \( x \) and \( y \) directions, respectively.
6. **Talbot images** - Calculate the Fresnel diffraction integral (5.56) for the case where the input field is taken to be given by \( u(x'',y'',z'' = 0^+) = u_0 t(x'',y'') \), where \( u_0 \) is a constant and the aperture transmission \( t(x'',y'') \) is given by

\[
t(x'',y'') = \frac{1 + \eta \cos \frac{2\pi x''}{a}}{2},
\]

where both \( \eta \) and \( a \) are positive constants.

7. Consider the aperture seen in Fig. 5.1, which contains a honeycomb array of transparent circular holes of radius \( R \). Outside the holes the aperture is opaque. The edge length of the hexagons is \( L \) (see Fig. 5.1). The aperture is positioned in the plane \( z = 0 \), and is normally illuminated by an incident monochromatic plane wave of wavelength \( \lambda \) and amplitude \( u_0 \). Assume that the total size of the aperture is much larger than \( L \), and that \( L \) is much larger than the radius of the holes \( R \). Employ the Fraunhofer diffraction integral to calculate the intensity \( I(x',y',z_0) \) on a screen positioned in the \( z = z_0 > 0 \) plane.

8. **Fresnel zone plate** - Consider an aperture having a transmission function \( t(x'',y'') \) given by

\[
t(x'',y'') = T_F \left( \sqrt{x''^2 + y''^2} \right),
\]

where the function \( T_F \) is given by
5.6. Solutions

\[ T_F (r) = \frac{1}{2} \left[ 1 + \text{sgn} \left( \cos (\gamma r^2) \right) \right] , \]  
(5.88)

the function \( \text{sgn} \) is the sign function, i.e.

\[ \text{sgn} (\theta) = \begin{cases} 
1 & \theta \geq 0 \\
-1 & \theta < 0 
\end{cases} , \]  
(5.89)

and \( \gamma \) is a positive constant. The aperture is positioned in the plane \( z = 0 \), and is normally illuminated by an incident monochromatic plane wave of wavelength \( \lambda \). Show that the aperture acts as a multi-focal lens.

a) Find the focal distance \( f_m \) of the \( m \)'th focal point.

b) Let \( P_m \) be the optical power that is delivered to the \( m \)'th focal point and let \( P_{in} \) be the optical power of the input plane wave. Calculate the relative power \( I_m = P_m / P_{in} \) that is delivered to the \( m \)'th focal point.

5.6 Solutions

1. The following holds [see Eq. (2.151)]

\[ \nabla \cdot (g \nabla u) = g \nabla^2 u + \nabla u \cdot \nabla g , \]  
(5.90)

\[ \nabla \cdot (u \nabla g) = u \nabla^2 g + \nabla u \cdot \nabla g , \]  
(5.91)

hence

\[ u \nabla^2 g - g \nabla^2 u = \nabla \cdot (u \nabla g - g \nabla u) . \]  
(5.92)

The above result together with the divergence theorem (2.68) lead to (5.82).

2. With the help of Eqs. (5.20) and (5.23) one finds that

\[ g (r) = - \frac{e^{ikr}}{4\pi r} , \]  
(5.93)

where \( r = |r| \). The following holds

\[ \nabla^2 g = \frac{1}{r^2} \frac{\partial}{\partial r} \left( r^2 \frac{\partial g}{\partial r} \right) = \frac{k^2 e^{ikr}}{4\pi r} , \]  
(5.94)

and thus for \( r \neq 0 \)

\[ (\nabla^2 + k^2) g (r) = 0 . \]  
(5.95)

Consider a sphere of radius \( r_0 \) centered at the origin. The integral of \( \nabla^2 g \) over the volume \( V \) of the sphere can be expressed using the divergence theorem (2.68) in terms of an an integral over the surface of the sphere \( S \).
Chapter 5. Scalar Diffraction Theory

\[ \int_V \nabla^2 g \, dv = \int_S \nabla g \cdot ds = \int_S e^{ikr} \frac{1 - ikr}{4\pi r^2} \hat{r} \cdot ds . \]  

(5.96)

The volume integral over \( k^2 g \) is given by

\[ \int_V k^2 g \, dv = -k^2 \int_0^{r_0} dr \, r^2 e^{i kr} . \]  

(5.97)

Thus, in the limit \( r_0 \to 0 \)

\[ \lim_{r_0 \to 0} \int_V (\nabla^2 + k^2) g \, dv = 1 , \]  

(5.98)

and hence Eq. (5.84) holds.

3. By employing the Green’s theorem (5.82) with the functions \( u(r) \) and \( g(r - r') \), which is given by Eq. (5.20), and by making use of Eqs. (5.1) and (5.84) one obtains

\[ \int_S \left( \frac{\partial g}{\partial n} - g \frac{\partial u}{\partial n} \right) ds = \int_V \left( u \nabla^2 g - g \nabla^2 u \right) dv \\
= \int_V \left( u \left( \nabla^2 + k^2 \right) g - g \left( \nabla^2 + k^2 \right) u \right) dv \\
= u(r) . \]  

(5.99)

The above result and the relation (5.23) lead to Eq. (5.85).

4. The scalar \( u(r'') \) in the plane \( z'' = 0^+ \) is taken to be given by

\[ u \left( x'', y'', z'' = 0^+ \right) = u_{inc} \left( x'', y'' \right) t \left( x'', y'' \right) , \]  

(5.100)

where for normal incident \( u_{inc} (x'', y'') \) is a constant denoted by \( u_0 \) and the aperture transmission is given by

\[ t (x'', y'') = \begin{cases} 1 & 0 < \rho'' \leq a \\ 0 & \rho'' > a \end{cases} , \]  

(5.101)

and

\[ \rho'' = \sqrt{x''^2 + y''^2} . \]  

(5.102)

In cylindrical coordinates

\[ x'' = \rho'' \cos \theta'' , \]  

(5.103)

\[ y'' = \rho'' \sin \theta'' , \]  

(5.104)

\[ x' = \rho' \cos \theta' , \]  

(5.105)

\[ y' = \rho' \sin \theta' , \]  

(5.106)

the Fresnel diffraction integral (5.56) becomes [see Eq. (5.30)]
5.6. Solutions

\[ u_1(r') = \frac{ik e^{ikz'}}{2\pi z'} \int_0^a \, d\rho'' e^{i\rho'' \cos \theta''} \, \int_0^\pi \, d\theta'' e^{ik \left( \rho'' \cos \theta'' \cos \theta' + \rho'' \sin \theta'' \sin \theta' \right)^2} \]

\[ = \frac{ik e^{ikz'} e^{ikz''}}{2\pi z'} \int_0^a \, d\rho'' e^{i\rho'' \cos \theta''} \, \int_0^\pi \, d\theta'' e^{ik \left( \rho'' \cos \theta'' \cos \theta' + \rho'' \sin \theta'' \sin \theta' \right)^2} \]

\[ = \frac{ik e^{ikz'} e^{ikz''}}{z'} \int_0^a \, d\rho'' e^{i\rho'' \cos \theta''} \left( J_0 \left( \frac{k \rho' \rho''}{z'} \right) \right) \]

\[ = \frac{ia}{\rho'} e^{ikz''} e^{ikz'} \left( J_1 \left( \frac{k a \rho'}{z'} \right) \right). \quad (5.107) \]

In the Fraunhofer (5.60) diffraction integral the terms \( e^{ika^2 s^2 / 2z'} \) is disregarded, and consequently \( u_1 (r') \) becomes

\[ u_1(r') = \frac{ika^2 u_0}{z'} e^{ikz'} \int_0^1 \, ds \, J_0 \left( \frac{k a \rho'}{z'} \right) \]

\[ = \frac{iau_0}{\rho'} e^{ikz'} \left( J_1 \left( \frac{k a \rho'}{z'} \right) \right). \quad (5.108) \]

5. The Fresnel diffraction integral (5.56) is given for the present case by

\[ u_1(r') = \frac{ik e^{ikz'} u_0}{2\pi z'} \int_{-a}^{-b} \, dx'' e^{ik (x'' - x')^2} \int_{-b}^{-a} \, dy'' e^{ik (y'' - y')^2} \]

\[ = \frac{ik e^{ikz'} u_0}{2\pi} \int_{a-}^{\alpha_+} \, dX e^{ikX^2} \int_{\beta_-}^{\beta_+} \, dY e^{ikY^2} \]

\[ = \frac{ik e^{ikz'} u_0}{2\pi} \left[ F^* (\alpha_+) - F^* (\alpha_-) \right] \left[ F^* (\beta_+) - F^* (\beta_-) \right]. \quad (5.109) \]

where the Fresnel function \( F (q) \) is given by

\[ F (q) = \int_0^q \, ds \, e^{-i q s^2}, \quad (5.110) \]

and where
\[ \alpha_\pm = \sqrt{\frac{k}{\pi z'}} (\pm a - x') , \]  
\[ \beta_\pm = \sqrt{\frac{k}{\pi z'}} (\pm b - x') . \] (5.111)

The Fraunhofer diffraction integral (5.60) is given for the present case by
\[ u_1 (r') = \frac{iku_0}{2\pi z'} e^{ikz'} \int_0^a dx'' e^{-ikz' x''} \int_{-b}^b dy'' e^{-ikz' y''} \]
\[ = \frac{2iku_0}{\pi z'} e^{ikz'} e^{i\frac{ka}{z'} x'} \text{sinc} \left( \frac{kax'}{z'} \right) \text{sinc} \left( \frac{kby'}{z'} \right). \] (5.112)

6. With the help of Eqs. (5.56) and (5.46) one obtains
\[ u_1 (r') = \frac{iku_0 e^{ikz'}}{2\pi z'} \int_{-\infty}^{\infty} dx'' \int_{-\infty}^{\infty} dy'' e^{ik (z' + \frac{x'^2 + y'^2}{2z'})} \]
\[ = -\frac{u_0 e^{ikz'}}{2} \left( 1 + \eta e^{-i\frac{ka}{z'} x'} \cos \frac{2\pi x'}{a} \right). \] (5.113)

Consider the case where the distance \( z' \) between the aperture and the screen is chosen such that the condition
\[ e^{-i\frac{ka}{z'} x'} = 1 \] (5.114)
is satisfied. For this case one finds with the help of Eq. (5.114) that \( |u_1 (r')|^2 = |u_0 t (x'', y'')|^2 \), i.e. an image of the grating is generated on the screen for such values of \( z' \).

7. It is convenient to introduce the so-called primitive vectors \( a_1 \) and \( a_2 \), which are taken to be given by
\[ a_1 = L \sqrt{3} \hat{y}, \] (5.115)
\[ a_2 = L \left( \frac{3}{2} \hat{x} + \sqrt{\frac{3}{2}} \hat{y} \right), \] (5.116)
and the so-called basis vectors \( c_1 \) and \( c_2 \), which are given by
\[ c_1 = 0, \] (5.117)
\[ c_2 = L \hat{x}, \] (5.118)
5.6. Solutions

Fig. 5.2. The primitive vectors $a_1$ and $a_2$.

In order to specify the locations of the centers of the holes in the array, which are given by

$$\rho_{n_1,n_2,m} = n_1 a_1 + n_2 a_2 + c_m,$$

(5.120)

where both $n_1$ and $n_2$ are integers, and where $m = 1$ ($m = 2$) for the red-colored (blue-colored) holes seen in Fig. 5.2. In the limit $R/L \to 0$ the holes are treated as point sources. The amplitude in the plane $z = 0^+$ is thus given by

$$u(x'', y'', z = 0^+) = u_0 \pi R^2 \sum_{n_1,n_2} \sum_{m=1,2} \delta (\rho'' - \rho_{n_1,n_2,m}),$$

(5.121)

where $\rho'' = (x'', y'')$. In the Fraunhofer approximation (5.60) the intensity on a screen is given by

$$I(x', y', z_0) = \left( \frac{\pi R^2}{\lambda z_0} \right)^2 |u_0|^2 \left| A \left( \frac{x'}{\lambda z_0}, \frac{y'}{\lambda z_0} \right) \right|^2,$$

(5.122)

where

$$A(K_x,K_y) = \sum_{n_1,n_2} \sum_{m=1,2} e^{-2\pi i K \cdot \rho_{n_1,n_2,m}}$$

(5.123)

and where $K = K_x \hat{x} + K_y \hat{y}$. Consider the so-called reciprocal lattice vectors $b_1$ and $b_2$, which are given by

$$b_1 = \frac{a_2 \times \hat{z}}{a_1 \cdot (a_2 \times \hat{z})} = \frac{-\frac{1}{2} \hat{x} + \frac{1}{\sqrt{3}} \hat{y}}{L},$$

(5.124)

$$b_2 = \frac{\hat{z} \times a_1}{a_1 \cdot (a_2 \times \hat{z})} = \frac{2 \hat{x}}{3L},$$

(5.125)
Chapter 5. Scalar Diffraction Theory

and which satisfy the following relation
\[ \mathbf{a}_n \cdot \mathbf{b}_m = \delta_{n,m}, \quad (5.126) \]

where \( n, m \in \{1, 2\} \). By expanding an arbitrary vector \( \mathbf{K} \) using the reciprocal lattice vectors \( \mathbf{b}_1 \) and \( \mathbf{b}_2 \) as
\[ \mathbf{K} = K_1 \mathbf{b}_1 + K_2 \mathbf{b}_2, \quad (5.127) \]

one obtains [see Eqs. (5.120), (5.123) and (5.126)]
\[ \mathcal{A}(K_x, K_y) = \sum_{n_1, n_2} \sum_{m=1,2} e^{-2\pi i (K_1 a_1 + K_2 a_2) \cdot (n_1 a_1 + n_2 a_2 + c_m)} = T(K) \sum_{n_1} e^{-2\pi i K_1 n_1} \sum_{n_2} e^{-2\pi i K_2 n_2}, \quad (5.128) \]

where [see Eqs. (5.118) and (5.119)]
\[ T(K) = \sum_{m=1,2} e^{-2\pi i K \cdot c_m} = 1 + e^{-2\pi i L \hat{x}}, \]
or [see Eqs. (5.124) and (5.125)]
\[ T(K) = 1 + e^{2\pi i (K_1 - 2K_2)}, \quad (5.129) \]

and thus
\[ |T(K)|^2 = 4 \cos^2 \left( \frac{\pi (K_1 - 2K_2)}{3} \right), \quad (5.130) \]

For an array of \( N \times N \) periods one finds with the help of the identity
\[ \sum_{n=-\frac{N}{2}}^{\frac{N}{2}} e^{-i\theta n} = \frac{\sin \left( \frac{N+1}{2} \theta \right)}{\sin \left( \frac{\theta}{2} \right)}, \quad (5.131) \]

that
\[ \mathcal{A}(K_x, K_y) = T(K) S(K_1) S(K_2), \quad (5.132) \]

where
\[ S(K) = \frac{\sin \left( (N + 1) \pi K \right)}{\sin \left( \pi K \right)}. \quad (5.133) \]

For \( N \gg 1 \) the function \( S^2(K) \) has sharp peaks near integer values of \( K \). Thus the intensity on the screen is expected to be high near the points \( \rho' \equiv (x',y') = \rho_{K_1,K_2} \), where
5.6. Solutions

Fig. 5.3. Intensity on the screen for an aperture made of a honeycomb array of holes.

\[ \rho_{K_1,K_2} = \lambda z_0 (K_1 b_1 + K_2 b_2) , \]  

and both \( K_1 \) and \( K_2 \) are integers. With the help of Eq. (5.130) one finds that for integer values of both \( K_1 \) and \( K_2 \) the term \( |T(K)|^2 \) is given by

\[ |T(K)|^2 = \begin{cases} 
4 & \text{if } \text{mod} (K_1 - 2K_2, 3) = 0 \\
1 & \text{else}
\end{cases} . \]  

The intensity on the screen is depicted by Fig. 5.3. The larger spots indicate the points \( \rho_{K_1,K_2} \) for which \( \text{mod} (K_1 - 2K_2, 3) = 0 \) and \( |T(K)|^2 = 4 \), whereas \( |T(K)|^2 = 1 \) for the other points.

8. Consider the function

\[ g(s) = \frac{1}{2} \left[ 1 + \text{sgn} (\cos(s)) \right] . \]  

The following holds \( g(s + 2\pi) = g(s) \), i.e. \( g(s) \) is periodic, and thus it can be Fourier expanded as
\[ g(s) = \sum_{m=-\infty}^{\infty} \left( \frac{1}{2\pi} \int_{-\pi}^{\pi} ds' g(s') e^{-ims} \right) e^{ims} \]
\[ = \sum_{m=-\infty}^{\infty} \left( \frac{1}{2\pi} \int_{-\pi}^{\pi} ds' e^{-ims} \right) e^{ims} \]
\[ = \sum_{m=-\infty}^{\infty} \frac{1}{\pi m} \sin \frac{\pi m}{2} e^{ims}. \]  
(5.137)

With the help of the above result (5.137) one finds that the amplitude transmission function (5.87) can be expanded as

\[ t(x'', y'') = \sum_{m=-\infty}^{\infty} \frac{1}{\pi m} \sin \frac{\pi m}{2} e^{im\gamma r'^2} \]
\[ = \sum_{m=-\infty}^{\infty} I_m^{1/2} e^{-i\pi r'^2 \over 2m}, \]  
(5.138)

where \( f_m \), which is given by

\[ f_m = -\frac{\pi}{m\lambda\gamma}, \]  
(5.139)

represents the \( m \)'th focal length [see Eq. (5.63) and recall that \( k = 2\pi/\lambda \)], the variable \( I_m \), which is given by

\[ I_m = \frac{\sin^2 \frac{\pi m}{\pi}}{\pi^2 m^2} = \begin{cases} \frac{1}{\pi m} & m = 0 \\ \frac{1}{(\pi m)^2} & m \text{ odd} \\ 0 & m \text{ even} \end{cases}, \]  
(5.140)

represents relative power that is delivered to the \( m \)'th focal point, and where

\[ r'^2 = x'^2 + y'^2. \]  
(5.141)

Note that for negative values of \( m \) the aperture acts as a diverging lens.
References

1. Ref
<table>
<thead>
<tr>
<th>Term</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABCD law, 126</td>
<td></td>
</tr>
<tr>
<td>ABCD ray matrix, 117</td>
<td></td>
</tr>
<tr>
<td>binormal, 75</td>
<td></td>
</tr>
<tr>
<td>Brewster’s angle, 60</td>
<td></td>
</tr>
<tr>
<td>Compton effect, 22</td>
<td></td>
</tr>
<tr>
<td>conductivity, 44</td>
<td></td>
</tr>
<tr>
<td>continuity equation, 18</td>
<td></td>
</tr>
<tr>
<td>Coulomb gauge, 42</td>
<td></td>
</tr>
<tr>
<td>curvature, 75</td>
<td></td>
</tr>
<tr>
<td>cylindrical coordinates, 82</td>
<td></td>
</tr>
<tr>
<td>D’Alembertian, 42</td>
<td></td>
</tr>
<tr>
<td>diffraction, 139</td>
<td></td>
</tr>
<tr>
<td>Doppler effect, 22</td>
<td></td>
</tr>
<tr>
<td>Drude model, 54</td>
<td></td>
</tr>
<tr>
<td>eikonal, 71</td>
<td></td>
</tr>
<tr>
<td>eikonal equation, 74</td>
<td></td>
</tr>
<tr>
<td>eikonal function, 96</td>
<td></td>
</tr>
<tr>
<td>electric displacement, 36</td>
<td></td>
</tr>
<tr>
<td>electric susceptibility, 44</td>
<td></td>
</tr>
<tr>
<td>energy-momentum four-vector, 12</td>
<td></td>
</tr>
<tr>
<td>fiber Bragg grating, 128</td>
<td></td>
</tr>
<tr>
<td>Fraunhofer diffraction integral, 147</td>
<td></td>
</tr>
<tr>
<td>Fresnel diffraction integral, 147</td>
<td></td>
</tr>
<tr>
<td>Fresnel equations, 54</td>
<td></td>
</tr>
<tr>
<td>gauge transformation, 41</td>
<td></td>
</tr>
<tr>
<td>gaussian beam, 126</td>
<td></td>
</tr>
<tr>
<td>graded index, 124</td>
<td></td>
</tr>
<tr>
<td>Green’s theorem, 151</td>
<td></td>
</tr>
<tr>
<td>Helmholtz equation, 50</td>
<td></td>
</tr>
<tr>
<td>Kirchhoff diffraction integral, 144</td>
<td></td>
</tr>
<tr>
<td>Leinard–Wiechert potential, 55</td>
<td></td>
</tr>
<tr>
<td>length contraction, 9</td>
<td></td>
</tr>
<tr>
<td>lensmaker’s equation, 119</td>
<td></td>
</tr>
<tr>
<td>Lifshitz formula, 54</td>
<td></td>
</tr>
<tr>
<td>light-like, 6</td>
<td></td>
</tr>
<tr>
<td>Lorentz transformation, 7</td>
<td></td>
</tr>
<tr>
<td>Lorenz condition, 42</td>
<td></td>
</tr>
<tr>
<td>Lorenz gauge, 42</td>
<td></td>
</tr>
<tr>
<td>Luneberg-Kline expansion, 73</td>
<td></td>
</tr>
<tr>
<td>Möbius transformation, 122</td>
<td></td>
</tr>
<tr>
<td>magnetic induction, 36</td>
<td></td>
</tr>
<tr>
<td>magnetic susceptibility, 44</td>
<td></td>
</tr>
<tr>
<td>Maxwell’s equations, 18, 35</td>
<td></td>
</tr>
<tr>
<td>Maxwell-Minkowski equations, 45</td>
<td></td>
</tr>
<tr>
<td>Minkowski metric, 5</td>
<td></td>
</tr>
<tr>
<td>paraxial approximation, 117</td>
<td></td>
</tr>
<tr>
<td>permeability, 44</td>
<td></td>
</tr>
<tr>
<td>permittivity, 44</td>
<td></td>
</tr>
<tr>
<td>plasma frequency, 55, 58</td>
<td></td>
</tr>
<tr>
<td>Poynting vector, 45</td>
<td></td>
</tr>
<tr>
<td>principle of least action, 99</td>
<td></td>
</tr>
<tr>
<td>proper time, 5</td>
<td></td>
</tr>
<tr>
<td>rapidity, 10</td>
<td></td>
</tr>
<tr>
<td>Rayleigh-Sommerfeld diffraction integral, 144</td>
<td></td>
</tr>
<tr>
<td>refraction index, 50</td>
<td></td>
</tr>
<tr>
<td>scattering time, 54</td>
<td></td>
</tr>
<tr>
<td>Snell’s law, 59</td>
<td></td>
</tr>
<tr>
<td>space-like, 6</td>
<td></td>
</tr>
<tr>
<td>space-time, 4</td>
<td></td>
</tr>
<tr>
<td>stationary phase approximation, 145</td>
<td></td>
</tr>
<tr>
<td>Talbot images, 152</td>
<td></td>
</tr>
<tr>
<td>time dilation, 9</td>
<td></td>
</tr>
<tr>
<td>time-like, 6</td>
<td></td>
</tr>
<tr>
<td>torsion, 76</td>
<td></td>
</tr>
<tr>
<td>total internal reflection, 60</td>
<td></td>
</tr>
<tr>
<td>Unruh-Davies effect, 23</td>
<td></td>
</tr>
<tr>
<td>Weyl’s expansion, 142</td>
<td></td>
</tr>
</tbody>
</table>