Eyal Buks

Wave Phenomena - Lecture Notes

September 22, 2021

Technion

Contents

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

$$\mathbf{F} = q\mathbf{E} , \tag{1.1}$$

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

$$\mathbf{E} = -\boldsymbol{\nabla}\phi , \tag{1.2}$$

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

$$\phi(\mathbf{r}) = \int \mathrm{d}^3\mathbf{r}' \frac{\rho(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} . \tag{1.3}$$

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

$$\boldsymbol{\nabla} \times \mathbf{E} = 0 . \tag{1.4}$$

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

$$\boldsymbol{\nabla}^2\phi = -\boldsymbol{\nabla} \cdot \mathbf{E} = -4\pi\rho . \tag{1.5}$$

Solution 1.1.1. The Poisson equation is easily derived with the help of the general identity

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

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 $4\pi r_{\mathrm{s}}^2$).

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. (1.99) 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 X is given by

$$X = (x_0, x_1, x_2, x_3)^{\mathrm{T}} , \tag{1.8}$$

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 $c\boldsymbol{\beta}$ with respect to the frame S (i.e. the dimensionless 3-vector $\boldsymbol{\beta}$ is the relative velocity of S' with respect to S in units of c). Let $X' = (x_0', x_1', x_2', x_3')^{\mathrm{T}}$ be the 4-vector of the same event as measured in S'.

Consider a second event having coordinates $X + \mathrm{d}X$, where

$$\mathrm{d}X = (\mathrm{d}x_0, \mathrm{d}x_1, \mathrm{d}x_2, \mathrm{d}x_3)^{\mathrm{T}} \tag{1.10}$$

is considered as infinitesimally small. The transformed 4-vector $\mathrm{d}X'$ is given by

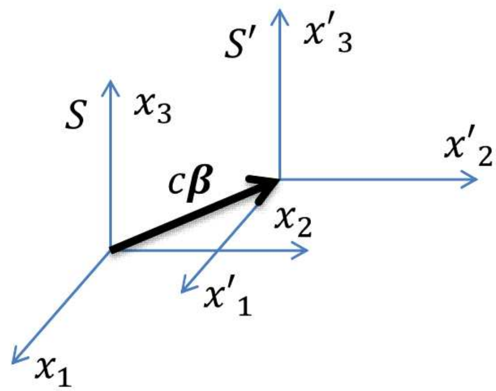

Fig. 1.1. The inertial frames of reference S and S'.

$$\mathrm{d}X' = \Lambda \mathrm{d}X \ , \tag{1.11}$$

where the 4×4 matrix Λ is given by

$$\Lambda = \frac{\partial (x'_0, x'_1, x'_2, x'_3)}{\partial (x_0, x_1, x_2, x_3)} \ . \tag{1.12}$$

Translational symmetry of space-time implies that Λ 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 $\mathrm{d}\tau$ corresponding to $\mathrm{d}X$ is defined by

$$(\mathrm{d}\tau)^2 \equiv c^{-2} (\mathrm{d}X)^{\mathrm{T}} \ \eta \ (\mathrm{d}X) \ , \tag{1.13}$$

where the so-called Minkowski metric η is given by

$$\eta = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \ , \tag{1.14}$$

thus

$$\begin{aligned} (\mathrm{d}\tau)^2 &= \frac{(\mathrm{d}x_0)^2 - (\mathrm{d}x_1)^2 - (\mathrm{d}x_2)^2 - (\mathrm{d}x_3)^2}{c^2} \\ &= (\mathrm{d}t)^2 \left(1 - \frac{\mathbf{u} \cdot \mathbf{u}}{c^2}\right) \ , \end{aligned} \tag{1.15}$$

where $\mathrm{d}t = c^{-1}\mathrm{d}x_0$, the velocity 3-vector $\mathbf{u}$ is given by

$$\mathbf{u} = \frac{\mathrm{d}\mathbf{r}}{\mathrm{d}t} = (u_1, u_2, u_3) = \left(\frac{\mathrm{d}x_1}{\mathrm{d}t}, \frac{\mathrm{d}x_2}{\mathrm{d}t}, \frac{\mathrm{d}x_3}{\mathrm{d}t}\right) , \tag{1.16}$$

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

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

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

Proof. As can be seen from Eq. (1.13), $(\mathrm{d}\tau)^2$ is independent on the direction of the velocity 3-vector $\mathbf{u}$. This property implies that the value $(\mathrm{d}\tau')^2$ 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 $(\mathrm{d}\tau)^2$) is expected to be independent on the direction of the 3-vector $\mathbf{v}'$, i.e. $(\mathrm{d}\tau')^2$ can be expressed as a function of $(\mathrm{d}\tau)^2$ and $v' = |\mathbf{v}'|$. Since the proper time is defined as infinitesimally small this function can be expressed as a linear function of $(\mathrm{d}\tau)^2$

$$(\mathrm{d}\tau')^2 = a_0(v') + a_1(v')(\mathrm{d}\tau)^2 , \tag{1.17}$$

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 $(\mathrm{d}\tau')^2 = 0$ when $(\mathrm{d}\tau)^2 = 0$, 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

$$(\mathrm{d}\tau'')^2 = a_1(v'')(\mathrm{d}\tau)^2 = a_1(v_{\mathrm{R}})\,a_1(v')(\mathrm{d}\tau)^2 , \tag{1.18}$$

where $v_{\mathrm{R}} = |\mathbf{v}_{\mathrm{R}}|$, and where $\mathbf{v}_{\mathrm{R}}$ is the relative velocity of frame S'' with respect to frame S', thus the following is required to hold

$$a_1(v'') = a_1(v_{\mathrm{R}})\,a_1(v') . \tag{1.19}$$

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)]

$$\Lambda^{\mathrm{T}}\eta\Lambda = \eta\ . \tag{1.20}$$

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

Exercise 1.2.1. Show that $\det \Lambda = \pm 1$

Solution 1.2.1. In general, $\det A^{\mathrm{T}} = \det A$ 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)] $(\det \Lambda)^2 = 1$.

Exercise 1.2.2. Find an expression for Λ^{-1}.

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

$$\Lambda^{-1} = \eta\Lambda^{\mathrm{T}}\eta\ . \tag{1.21}$$

Exercise 1.2.3. Show that if both Λ_1 and Λ_2 are Lorentz transformations then $\Lambda_1\Lambda_2$ is a Lorentz transformation.

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

$$(\Lambda_1\Lambda_2)^{\mathrm{T}}\,\eta\Lambda_1\Lambda_2 = \Lambda_2^{\mathrm{T}}\Lambda_1^{\mathrm{T}}\eta\Lambda_1\Lambda_2 = \Lambda_2^{\mathrm{T}}\eta\Lambda_2 = \eta\ , \tag{1.22}$$

and thus $\Lambda_1\Lambda_2$ is a Lorentz transformation.

As an example, consider the case where the inertial frame of reference S' moves at a constant velocity βc 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 Λ, which relates the vectors of coordinates $\mathrm{d}X$ and $\mathrm{d}X'$ by $\mathrm{d}X' = \Lambda\mathrm{d}X$ [see Eq. (1.11)], can be expressed in a block form as

$$\Lambda = \left(\begin{array}{c|c} B_1 & 0 \\ \hline 0 & \begin{matrix} 1 & 0 \\ 0 & 1 \end{matrix} \end{array}\right)\ , \tag{1.23}$$

where

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

is a 2×2 matrix relating the time and x_1 coordinates

$$\begin{pmatrix} c\mathrm{d}t' \\ \mathrm{d}x_1' \end{pmatrix} = B_1 \begin{pmatrix} c\mathrm{d}t \\ \mathrm{d}x_1 \end{pmatrix}\ . \tag{1.25}$$

Claim. The 2×2 matrix B_1 is given by

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

where

$$\gamma = \frac{1}{\sqrt{1-\beta^2}} . \tag{1.27}$$

Proof. For events occurring at $\mathrm{d}x_1' = 0$ the following holds (since the relative velocity of S' with respect to S is βc)

$$\frac{1}{c}\frac{\mathrm{d}x_1}{\mathrm{d}t} = \beta , \tag{1.28}$$

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

$$0 = \mathrm{d}x_1' = b_{21} c\mathrm{d}t + b_{22}\mathrm{d}x_1 = \left(\frac{b_{21}}{\beta} + b_{22}\right)\mathrm{d}x_1 . \tag{1.29}$$

Similarly, for events occurring at $\mathrm{d}x_1 = 0$ the following holds (since the relative velocity of S with respect to S' is $-\beta c$)

$$\frac{1}{c}\frac{\mathrm{d}x_1'}{\mathrm{d}t'} = -\beta , \tag{1.30}$$

and thus [see of Eq. (1.25)]

$$\frac{1}{c}\frac{\mathrm{d}x_1'}{\mathrm{d}t'} = \frac{b_{21}}{b_{11}} = -\beta , \tag{1.31}$$

and therefore B_1 can be expressed as

$$B_1 = \gamma \begin{pmatrix} 1 & \frac{b_{12}}{\gamma} \\ -\beta & 1 \end{pmatrix} , \tag{1.32}$$

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

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

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 $\mathrm{d}x_1 = 0$. For this case $\mathrm{d}t = \mathrm{d}\tau$, i.e. the time difference between the two events $\mathrm{d}t$ as measured in a frame S at which both events occur at the same location is the proper time $\mathrm{d}\tau$ [see Eq. (1.15)]. With the help of Eq. (1.25) one finds that
$$\mathrm{d}t' = \gamma \mathrm{d}\tau \geq \mathrm{d}\tau \ . \tag{1.34}$$
The time dilation factor γ 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 $\mathrm{d}x_1$. Let $\mathrm{d}x_1'$ 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 $\mathrm{d}x_1'$ in the frame S' is associated with two events having the same time, i.e. $\mathrm{d}t' = 0$. With the help of Eq. (1.25) one finds that
$$\begin{aligned}\begin{pmatrix} c\mathrm{d}t \\ \mathrm{d}x_1 \end{pmatrix} &= B_1^{-1} \begin{pmatrix} 0 \\ \mathrm{d}x_1' \end{pmatrix} \\ &= \gamma \begin{pmatrix} 1 & \beta \\ \beta & 1 \end{pmatrix} \begin{pmatrix} 0 \\ \mathrm{d}x_1' \end{pmatrix} \ ,\end{aligned} \tag{1.36}$$
and thus
$$\mathrm{d}x_1' = \frac{\mathrm{d}x_1}{\gamma} \leq \mathrm{d}x_1 \ , \tag{1.37}$$
i.e. the length of the rod $\mathrm{d}x_1'$ 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 4×4 matrix $B(\boldsymbol{\beta})$ is defined by
$$B(\boldsymbol{\beta}) = \exp\left(-\kappa \hat{\boldsymbol{\beta}} \cdot \Sigma\right) \ , \tag{1.38}$$
where $\hat{\boldsymbol{\beta}}$ is a unit vector pointing in the direction of $\boldsymbol{\beta}$ (i.e. $\boldsymbol{\beta} = \beta\hat{\boldsymbol{\beta}}$ where $\beta = |\boldsymbol{\beta}|$), the components of the matrix vector $\Sigma = (\Sigma_1, \Sigma_3, \Sigma_3)$ are given by
$$\Sigma_1 = \begin{pmatrix} 0&1&0&0 \\ 1&0&0&0 \\ 0&0&0&0 \\ 0&0&0&0 \end{pmatrix} , \Sigma_2 = \begin{pmatrix} 0&0&1&0 \\ 0&0&0&0 \\ 1&0&0&0 \\ 0&0&0&0 \end{pmatrix} , \Sigma_3 = \begin{pmatrix} 0&0&0&1 \\ 0&0&0&0 \\ 0&0&0&0 \\ 1&0&0&0 \end{pmatrix} , \tag{1.39}$$

and where the so-called rapidity κ is given by

$$\kappa = \tanh^{-1} \beta \ . \tag{1.40}$$

Show that

$$B(\boldsymbol{\beta}) = \begin{pmatrix} \gamma & -\gamma\beta_1 & -\gamma\beta_2 & -\gamma\beta_3 \\ -\gamma\beta_1 & 1+\frac{(\gamma-1)\beta_1^2}{\beta^2} & \frac{(\gamma-1)\beta_1\beta_2}{\beta^2} & \frac{(\gamma-1)\beta_1\beta_3}{\beta^2} \\ -\gamma\beta_2 & \frac{(\gamma-1)\beta_2\beta_1}{\beta^2} & 1+\frac{(\gamma-1)\beta_2^2}{\beta^2} & \frac{(\gamma-1)\beta_2\beta_3}{\beta^2} \\ -\gamma\beta_3 & \frac{(\gamma-1)\beta_3\beta_1}{\beta^2} & \frac{(\gamma-1)\beta_3\beta_2}{\beta^2} & 1+\frac{(\gamma-1)\beta_3^2}{\beta^2} \end{pmatrix} , \tag{1.41}$$

where

$$\gamma = \frac{1}{\sqrt{1-\beta^2}} \ . \tag{1.42}$$

Solution 1.2.4. The following holds

$$\left(\hat{\boldsymbol{\beta}} \cdot \Sigma\right)^n = \begin{cases} B_{\mathrm{O}}\left(\hat{\boldsymbol{\beta}}\right) & n = 1, 3, \cdots \\ B_{\mathrm{E}}\left(\hat{\boldsymbol{\beta}}\right) & n = 2, 4, \cdots \end{cases} , \tag{1.43}$$

where

$$B_{\mathrm{O}}\left(\hat{\boldsymbol{\beta}}\right) = \begin{pmatrix} 0 & b_1 & b_2 & b_3 \\ b_1 & 0 & 0 & 0 \\ b_2 & 0 & 0 & 0 \\ b_3 & 0 & 0 & 0 \end{pmatrix} , \tag{1.44}$$

$$B_{\mathrm{E}}\left(\hat{\boldsymbol{\beta}}\right) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & b_1^2 & b_1 b_2 & b_1 b_3 \\ 0 & b_2 b_1 & b_2^2 & b_2 b_3 \\ 0 & b_3 b_1 & b_3 b_2 & b_3^2 \end{pmatrix} , \tag{1.45}$$

and where

$$\hat{\boldsymbol{\beta}} = (b_1, b_2, b_3) = \frac{1}{\beta}(\beta_1, \beta_2, \beta_3) \ . \tag{1.46}$$

This result together with Eq. (1.38) leads to

$$\begin{aligned} B(\boldsymbol{\beta}) &= \exp\left(-\kappa\hat{\boldsymbol{\beta}} \cdot \Sigma\right) \\ &= -\sinh\left(\kappa\hat{\boldsymbol{\beta}} \cdot \Sigma\right) + \cosh\left(\kappa\hat{\boldsymbol{\beta}} \cdot \Sigma\right) \\ &= \mathbf{1} - \sinh(\kappa) B_{\mathrm{O}}\left(\hat{\boldsymbol{\beta}}\right) + (\cosh(\kappa) - 1) B_{\mathrm{E}}\left(\hat{\boldsymbol{\beta}}\right) , \end{aligned} \tag{1.47}$$

where $\mathbf{1}$ is the identity matrix and where [see Eq. (1.40)]

$$\sinh(\kappa) = \beta\gamma \; , \tag{1.48}$$
$$\cosh(\kappa) = \gamma \; . \tag{1.49}$$

Combining the above results lead to Eq. (1.41). Note that the matrix $B(\boldsymbol{\beta})$ (1.41) satisfies the condition (1.20), i.e. $B(\boldsymbol{\beta})^{\mathrm{T}} \eta B(\boldsymbol{\beta}) = \eta$, and thus $B(\boldsymbol{\beta})$ is a Lorentz transformation.

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

Solution 1.2.5. The coordinates of frame S are chosen such that the velocity $\mathbf{u}$ is pointing in the x_1 direction. For this case the Lorentz transformation law of the space-time 4-vector $(c\mathrm{d}t, \mathrm{d}x_1, \mathrm{d}x_2, \mathrm{d}x_3)^{\mathrm{T}}$ reads [see Eq. (1.26)]

$$\begin{pmatrix} c\mathrm{d}t' \\ \mathrm{d}x_1' \\ \mathrm{d}x_2' \\ \mathrm{d}x_3' \end{pmatrix} = \begin{pmatrix} \gamma & -\gamma\beta & 0 & 0 \\ -\gamma\beta & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} c\mathrm{d}t \\ \mathrm{d}x_1 \\ \mathrm{d}x_2 \\ \mathrm{d}x_3 \end{pmatrix} , \tag{1.50}$$

or

$$\mathrm{d}t' = \gamma \mathrm{d}t - \frac{\gamma\beta}{c}\mathrm{d}x_1 \; , \tag{1.51}$$
$$\mathrm{d}x_1' = \gamma \mathrm{d}x_1 - \gamma\beta c\mathrm{d}t \; , \tag{1.52}$$
$$\mathrm{d}x_2' = \mathrm{d}x_2 \; , \tag{1.53}$$
$$\mathrm{d}x_3' = \mathrm{d}x_3 \; , \tag{1.54}$$

where $\beta = u/c$ and $\gamma = 1/\sqrt{1-\beta^2}$. Dividing Eqs. (1.52), (1.53), and (1.54) by $\mathrm{d}t'$ yields

$$v_1' = \frac{v_1 - \beta c}{1 - \frac{\beta v_1}{c}} \; , \tag{1.55}$$
$$v_2' = \frac{v_2}{\gamma\left(1 - \frac{\beta v_1}{c}\right)} \; , \tag{1.56}$$
$$v_3' = \frac{v_3}{\gamma\left(1 - \frac{\beta v_1}{c}\right)} \; , \tag{1.57}$$

where

$$v_n = \frac{\mathrm{d}x_n}{\mathrm{d}t} \; , v_n' = \frac{\mathrm{d}x_n'}{\mathrm{d}t'} \; , \tag{1.58}$$

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

$$(v_1', v_2', v_3') = \frac{\left(\frac{v_1}{\gamma}, \frac{v_2}{\gamma}, \frac{v_3}{\gamma}\right) - \left(u - \left(1 - \frac{1}{\gamma}\right) v_1, 0, 0\right)}{1 - \frac{\mathbf{u}\cdot\mathbf{v}}{c^2}} \; . \tag{1.59}$$

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

$$\mathbf{v}' = \frac{\frac{\mathbf{v}}{\gamma} - \left[1 - \left(1 - \frac{1}{\gamma}\right)\frac{\mathbf{u}\cdot\mathbf{v}}{u^2}\right]\mathbf{u}}{1 - \frac{\mathbf{u}\cdot\mathbf{v}}{c^2}} \ . \tag{1.60}$$

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

$$\hat{\mathbf{n}}' = \frac{\hat{\mathbf{n}} - \gamma\left(1 - \frac{\gamma}{1+\gamma}\frac{\mathbf{u}\cdot\hat{\mathbf{n}}}{c}\right)\frac{\mathbf{u}}{c}}{\gamma\left(1 - \frac{\mathbf{u}\cdot\hat{\mathbf{n}}}{c}\right)} \ , \tag{1.61}$$

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

$$\cos\theta' = \frac{\cos\theta - \beta}{1 - \beta\cos\theta} \ , \tag{1.62}$$

where $\mathbf{u}\cdot\hat{\mathbf{n}} = \beta c\cos\theta$ and $\mathbf{u}\cdot\hat{\mathbf{n}}' = \beta c\cos\theta'$ [note that $\beta^2\gamma^2/\left(1+\gamma\right) = \gamma - 1$].

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), $\mathrm{d}\tau = \mathrm{d}t$ in that frame, which implies that the proper time $\mathrm{d}\tau$ is the time difference between the events as measured in S. The transformation of $\mathrm{d}X$ into a frame S' having relative velocity $\mathbf{u} = c\boldsymbol{\beta}$ with respect to S is given by [see Eq. (1.11)]

$$\mathrm{d}X' = \Lambda\mathrm{d}X \ , \tag{1.63}$$

where $\Lambda = B\left(\boldsymbol{\beta}\right)$ [see Eq. (1.38)].

The Energy-Momentum 4-Vector. The energy-momentum 4-vector P is defined by

$$P = m\frac{\mathrm{d}X}{\mathrm{d}\tau} \ . \tag{1.64}$$

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

$$P = \left(mc, 0, 0, 0\right)^{\mathrm{T}} \ . \tag{1.65}$$

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 $\mathrm{d}\tau$ and Eq. (1.11) together imply that P is Lorentz transformed according to

$$P' = \Lambda P \; . \tag{1.66}$$

In the frame S' it is given by

$$P' = m\left(\frac{\mathrm{d}x'_0}{\mathrm{d}\tau}, \frac{\mathrm{d}x'_1}{\mathrm{d}\tau}, \frac{\mathrm{d}x'_2}{\mathrm{d}\tau}, \frac{\mathrm{d}x'_3}{\mathrm{d}\tau}\right)^{\mathrm{T}} \; . \tag{1.67}$$

With the help of Eq. (1.34) one finds that

$$\frac{1}{\mathrm{d}\tau} = \frac{\mathrm{d}t'}{\mathrm{d}\tau}\frac{1}{\mathrm{d}t'} = \gamma\frac{1}{\mathrm{d}t'} \; , \tag{1.68}$$

where

$$\gamma = \frac{1}{\sqrt{1-\frac{u^2}{c^2}}} \; , \tag{1.69}$$

and where $u = |\mathbf{u}| = c\,|\boldsymbol{\beta}|$, and thus P' can be expressed as

$$P' = \left(p'_0, p'_1, p'_2, p'_3\right)^{\mathrm{T}} = \left(\frac{E'}{c}, m\gamma u_1, m\gamma u_2, m\gamma u_3\right)^{\mathrm{T}} \; , \tag{1.70}$$

where

$$E' = mc^2\gamma \; . \tag{1.71}$$

Note that the following holds [see Eq. (1.69)]

$$E' = mc^2 + \frac{mu^2}{2} + O\left(\beta^4\right) \; , \tag{1.72}$$

$$\left(p'_1, p'_2, p'_3\right) = m\mathbf{u} + O\left(\beta^3\right) \; , \tag{1.73}$$

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 $\left(p'_1, p'_2, p'_3\right)$ becomes its Newtonian momentum vector.

The fact that P is transformed by a Lorentz transformation implies that the quantity $p_0^2 - p_1^2 - p_2^2 - p_3^2$ is invariant. With the help of Eqs. (1.65) and (1.70) one finds that

$$m^2c^2 = \frac{E^2}{c^2} - p^2 \; . \tag{1.74}$$

While the left hand side of (1.74) is frame independent, both energy E and momentum p of the particle are frame dependent. For massless particles Eq. (1.74) reads

$$E = cp \; . \tag{1.75}$$

Consider a scattering process involving N_{in} incoming particles and N_{out} outgunning particles. The energy-momentum conservation law implies that

$$\sum_{n=1}^{N_{\text{in}}} P_{n,\text{in}} = \sum_{n=1}^{N_{\text{out}}} P_{n,\text{out}} \; , \tag{1.76}$$

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. (1.70), the momentum of a particle that is conserved according to the theory of special relativity is given by $m\gamma\mathbf{u}$. 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-dependant mass $m\gamma$, where m is the rest mass.

The Force 4-Vector. The force 4-vector F is defined by

$$F = \frac{\mathrm{d}P}{\mathrm{d}\tau} \; . \tag{1.77}$$

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

$$F' = \Lambda F \; . \tag{1.78}$$

The force 4-vector F is related to the force 3-vector $\mathbf{f}$, which is defined by

$$\mathbf{f} = \frac{\mathrm{d}\mathbf{p}}{\mathrm{d}t} \; , \tag{1.79}$$

by [see Eqs. (1.34), (1.64) and (1.70)]

$$F = \left(\frac{1}{c}\frac{\mathrm{d}E}{\mathrm{d}\tau}, \frac{\mathrm{d}\mathbf{p}}{\mathrm{d}\tau} \right)^{\mathrm{T}} = \gamma \left(\frac{1}{c}\frac{\mathrm{d}E}{\mathrm{d}t}, \mathbf{f} \right)^{\mathrm{T}} \; . \tag{1.80}$$

Exercise 1.2.6. Show that

$$\frac{\mathrm{d}E}{\mathrm{d}t} = \mathbf{f} \cdot \mathbf{v} \; , \tag{1.81}$$

where $\mathbf{f}$ is the 3-vector force and $\mathbf{v}$ 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 $P^{\mathrm{T}}\eta F$, which is given by [see Eqs. (1.70) and (1.80)]

$$P^{\mathrm{T}}\eta F = m\gamma^2 \left(c \;\mathbf{v} \right) \eta \begin{pmatrix} \frac{1}{c}\frac{\mathrm{d}E}{\mathrm{d}t} \\ \mathbf{f} \end{pmatrix} = m\gamma^2 \left(\frac{\mathrm{d}E}{\mathrm{d}t} - \mathbf{f} \cdot \mathbf{v} \right) \; . \tag{1.82}$$

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. (1.71)] and $\mathbf{v}' = 0$, and thus $\left(P^{\mathrm{T}}\eta F\right)'$ vanishes (it is assumed that mass of the particle m is a constant). The fact that $P^{\mathrm{T}}\eta F$ is invariant under Lorentz transformation [see Eqs. (1.66) and (1.78)] implies that $P^{\mathrm{T}}\eta F = 0$, and thus (1.81) 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 $\mathbf{f}$ and $\mathbf{v}$, respectively. Calculate the 3-vector force $\mathbf{f}'$ as measured in a frame S' moving at a constant velocity $\mathbf{u}$ with respect to the frame S.

Solution 1.2.7. The coordinates of frame S are chosen such that the velocity $\mathbf{u}$ 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)]

$$\frac{\mathrm{d}E'}{\mathrm{d}t'} = \gamma \frac{\mathrm{d}t}{\mathrm{d}t'} \left(\frac{\mathrm{d}E}{\mathrm{d}t} - \beta c f_1 \right) , \tag{1.83}$$

$$f_1' = \gamma \frac{\mathrm{d}t}{\mathrm{d}t'} \left(f_1 - \frac{\beta}{c} \frac{\mathrm{d}E}{\mathrm{d}t} \right) , \tag{1.84}$$

$$f_2' = \frac{\mathrm{d}t}{\mathrm{d}t'} f_2 , \tag{1.85}$$

$$f_3' = \frac{\mathrm{d}t}{\mathrm{d}t'} f_3 . \tag{1.86}$$

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

$$\frac{\mathrm{d}t}{\mathrm{d}t'} = \frac{1}{\gamma \left(1 - \frac{\beta}{c} \frac{\mathrm{d}x_1}{\mathrm{d}t} \right)} = \frac{1}{\gamma \left(1 - \frac{\mathbf{u}\cdot\mathbf{v}}{c^2} \right)} , \tag{1.87}$$

and thus [see Eq. (1.81)]

$$f_1' = \frac{f_1 - \frac{\beta}{c} \mathbf{f}\cdot\mathbf{v}}{1 - \frac{\mathbf{u}\cdot\mathbf{v}}{c^2}} , \tag{1.88}$$

$$f_2' = \frac{f_2}{\gamma \left(1 - \frac{\mathbf{u}\cdot\mathbf{v}}{c^2} \right)} , \tag{1.89}$$

$$f_3' = \frac{f_3}{\gamma \left(1 - \frac{\mathbf{u}\cdot\mathbf{v}}{c^2} \right)} . \tag{1.90}$$

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

$$\mathbf{a} = \mathbf{a}_{\parallel} + \mathbf{a}_{\perp} , \tag{1.91}$$

where

$$\mathbf{a}_{\parallel} = \frac{(\mathbf{u}\cdot\mathbf{a})}{u^2} \mathbf{u} , \tag{1.92}$$

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

$$\mathbf{f} = \mathbf{f}'_\parallel + \gamma \mathbf{f}'_\perp + \gamma \frac{\mathbf{v} \times (\mathbf{u} \times \mathbf{f}'_\perp)}{c^2} \ . \tag{1.93}$$

Solution 1.2.8. In vectorial notation Eq. (1.88), (1.89) and (1.90) become

$$\mathbf{f}'_\parallel = \frac{\mathbf{f}_\parallel - \frac{(\mathbf{f}\cdot\mathbf{v})\mathbf{u}}{c^2}}{1 - \frac{\mathbf{u}\cdot\mathbf{v}}{c^2}} \ , \tag{1.94}$$

$$\mathbf{f}'_\perp = \frac{\mathbf{f}_\perp}{\gamma \left(1 - \frac{\mathbf{u}\cdot\mathbf{v}}{c^2}\right)} \ . \tag{1.95}$$

With the help of the vector identity

$$\mathbf{A} \times (\mathbf{B} \times \mathbf{C}) = (\mathbf{A} \cdot \mathbf{C})\, \mathbf{B} - (\mathbf{A} \cdot \mathbf{B})\, \mathbf{C} \ , \tag{1.96}$$

one finds that

$$(\mathbf{f} \cdot \mathbf{v})\, \mathbf{u} = \mathbf{v} \times (\mathbf{u} \times \mathbf{f}) + (\mathbf{v} \cdot \mathbf{u})\, \mathbf{f} \ , \tag{1.97}$$

and thus Eq. (1.94) can be rewritten as (note that $\mathbf{u} \times \mathbf{f} = \mathbf{u} \times \mathbf{f}_\perp$)

$$\begin{aligned}
\mathbf{f}'_\parallel &= \frac{\mathbf{f}_\parallel - \frac{\mathbf{v}\times(\mathbf{u}\times\mathbf{f}) + (\mathbf{v}\cdot\mathbf{u})\mathbf{f}}{c^2}}{1 - \frac{\mathbf{u}\cdot\mathbf{v}}{c^2}} \\
&= \mathbf{f}_\parallel - \frac{\frac{\mathbf{v}\times(\mathbf{u}\times\mathbf{f}_\perp)}{c^2} + \frac{(\mathbf{v}\cdot\mathbf{u})\mathbf{f}_\perp}{c^2}}{1 - \frac{\mathbf{u}\cdot\mathbf{v}}{c^2}} \\
&= \mathbf{f}_\parallel - \gamma c^{-2} \left[\mathbf{v} \times (\mathbf{u} \times \mathbf{f}'_\perp) + (\mathbf{v} \cdot \mathbf{u})\, \mathbf{f}'_\perp \right] \ .
\end{aligned} \tag{1.98}$$

The above result together with Eq. (1.95) lead to (1.93).

Exercise 1.2.9. Show that Eq. (1.93) can be rewritten as

$$\mathbf{f} = \mathbf{f}_\mathrm{E} + \frac{\mathbf{v} \times \mathbf{f}_\mathrm{B}}{c} \ , \tag{1.99}$$

where

$$\mathbf{f}_\mathrm{E} = \mathbf{f}'_\parallel + \gamma \mathbf{f}'_\perp \ , \tag{1.100}$$

and where

$$\mathbf{f}_\mathrm{B} = \frac{\mathbf{u} \times \mathbf{f}_\mathrm{E}}{c} \ . \tag{1.101}$$

Solution 1.2.9. On one hand Eqs. (1.93) and (1.100) yield

$$\mathbf{f} - \mathbf{f}_{\mathrm{E}} = \frac{\gamma \mathbf{v} \times (\mathbf{u} \times \mathbf{f}'_{\perp})}{c^2} \; . \tag{1.102}$$

On the other hand since $\mathbf{f}'_{\parallel}$ is parallel to $\mathbf{u}$ the following holds [see Eq. (1.100)]

$$\frac{\mathbf{u} \times \mathbf{f}_{\mathrm{E}}}{c} = \frac{\gamma \mathbf{u} \times \mathbf{f}'_{\perp}}{c} \; , \tag{1.103}$$

and thus Eq. (1.102) can be rewritten as

$$\mathbf{f} - \mathbf{f}_{\mathrm{E}} = \frac{\mathbf{v} \times \left(\frac{\mathbf{u} \times \mathbf{f}_{\mathrm{E}}}{c}\right)}{c} \; , \tag{1.104}$$

in agreement with Eq. (1.99).

1.3 Transformation of Electrostatic Force

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

$$\boldsymbol{\nabla}' \cdot \mathbf{E}' = 4\pi\rho' \; , \tag{1.105}$$

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

$$\boldsymbol{\nabla}' \times \mathbf{E}' = 0 \; . \tag{1.106}$$

The above-mentioned quantities that are labeled by a prime ($\mathbf{F}'$, ρ' and $\mathbf{E}'$) 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 $\mathbf{v}$. Let $\mathbf{u}$ be the relative velocity of frame S' with respect to S.

With the help of Eq. (1.99) one finds that the force on the test particle $\mathbf{F}$ as measured in frame S can be expressed as

$$\mathbf{F} = q\left(\mathbf{E} + \frac{\mathbf{v} \times \mathbf{B}}{c}\right) \; , \tag{1.107}$$

where $\mathbf{E}$ is given by

$$\mathbf{E} = \mathbf{E}'_{\parallel} + \gamma \mathbf{E}'_{\perp} \; , \tag{1.108}$$

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

$$\mathbf{B} = \frac{\mathbf{u} \times \mathbf{E}}{c} \; . \tag{1.109}$$

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

1.4 Charge and Current Density

Let ρ be the charge distribution and let $\mathbf{J}$ 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 $\mathbf{J}'$ as measured in S' vanishes. Consider an infinitesimal volume $\mathrm{d}V'$ containing N particle having charge q each. In the frame S the measured volume $\mathrm{d}V$ is smaller due to length contraction $\mathrm{d}V = \gamma^{-1}\mathrm{d}V'$ [see Eq. (1.37)], and consequently (note that both q and N are required to be frame independent)

$$\rho = \gamma\rho' . \tag{1.110}$$

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

$$\mathbf{J} = \gamma\rho'\mathbf{u} . \tag{1.111}$$

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

$$J = (c\rho, J_1, J_2, J_3)^{\mathrm{T}} , \tag{1.112}$$

is Lorentz transformed according to

$$J' = \Lambda J . \tag{1.113}$$

For an alternative definition of the current 4-vector see Eq. (1.146). Consider the quantity ∂J, where the 4-vector ∂ is defined by

$$\partial = \left(c^{-1}\frac{\partial}{\partial t}, \frac{\partial}{\partial x_1}, \frac{\partial}{\partial x_2}, \frac{\partial}{\partial x_3} \right) . \tag{1.114}$$

As can be seen from Eq. (1.12) ∂ is transformed according to

$$\partial' = \partial\Lambda^{-1} , \tag{1.115}$$

and thus ∂J, which is given by

$$\partial J = \frac{\partial\rho}{\partial t} + \boldsymbol{\nabla}\cdot\mathbf{J} , \tag{1.116}$$

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

$$0 = \frac{\partial\rho}{\partial t} + \boldsymbol{\nabla}\cdot\mathbf{J} , \tag{1.117}$$

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

1.5 Maxwell's Equations

Claim. The fields $\mathbf{E}$ (1.108) and $\mathbf{B}$ (1.109), which are used to express the force $\mathbf{F}$ acting on the test particle according to Eq. (1.107), satisfy the Maxwell's Eqs.

$$\boldsymbol{\nabla}\times\mathbf{B}=\frac{4\pi}{c}\mathbf{J}+\frac{1}{c}\frac{\partial\mathbf{E}}{\partial t}\ , \tag{1.118}$$

$$\boldsymbol{\nabla}\times\mathbf{E}=-\frac{1}{c}\frac{\partial\mathbf{B}}{\partial t}\ , \tag{1.119}$$

$$\boldsymbol{\nabla}\cdot\mathbf{E}=4\pi\rho\ , \tag{1.120}$$

$$\boldsymbol{\nabla}\cdot\mathbf{B}=0\ . \tag{1.121}$$

Proof. As can be seen from Eq. (1.115), the following holds [see Eq. (1.26)]

$$\frac{\partial}{\partial t}=\gamma\left(\frac{\partial}{\partial t'}-c\beta\frac{\partial}{\partial x_1'}\right)\ , \tag{1.122}$$

$$\frac{\partial}{\partial x_1}=\gamma\left(\frac{\partial}{\partial x_1'}-\frac{\beta}{c}\frac{\partial}{\partial t'}\right)\ , \tag{1.123}$$

$$\frac{\partial}{\partial x_2}=\frac{\partial}{\partial x_2'}\ , \tag{1.124}$$

$$\frac{\partial}{\partial x_3}=\frac{\partial}{\partial x_3'}\ . \tag{1.125}$$

Since $\mathbf{E}$ does not depend on t one finds that [see Eqs. (1.108), (1.123), (1.124) and (1.125)]

$$\begin{aligned}\boldsymbol{\nabla}\cdot\mathbf{E}&=\frac{\partial E_1}{\partial x_1}+\frac{\partial E_2}{\partial x_2}+\frac{\partial E_3}{\partial x_3}\\&=\frac{\gamma\partial E_1'}{\partial x_1'}+\frac{\gamma\partial E_2'}{\partial x_2'}+\frac{\gamma\partial E_3'}{\partial x_3'}\\&=\gamma\boldsymbol{\nabla}'\cdot\mathbf{E}'\ .\end{aligned} \tag{1.126}$$

The above result together with Eqs. (1.105) and (1.110) lead to Eq. (1.120)

$$\boldsymbol{\nabla}\cdot\mathbf{E}=4\pi\gamma\rho'=4\pi\rho\ . \tag{1.127}$$

Using the vector identity

$$\nabla\times(\mathbf{A}\times\mathbf{B})=\mathbf{A}\left(\nabla\cdot\mathbf{B}\right)-\mathbf{B}\left(\nabla\cdot\mathbf{A}\right)+\left(\mathbf{B}\cdot\nabla\right)\mathbf{A}-\left(\mathbf{A}\cdot\nabla\right)\mathbf{B}\ , \tag{1.128}$$

one obtains (recall that $\mathbf{u}$ is a constant vector)

$$\boldsymbol{\nabla}\times(\mathbf{u}\times\mathbf{E})=\mathbf{u}\left(\nabla\cdot\mathbf{E}\right)-\left(\mathbf{u}\cdot\nabla\right)\mathbf{E}\ , \tag{1.129}$$

and thus [see Eqs. (1.109), (1.123) and (1.127)]

$$
\begin{aligned}
\boldsymbol{\nabla} \times \mathbf{B} &= \frac{\mathbf{u}\,(\nabla \cdot \mathbf{E}) - (\mathbf{u} \cdot \nabla)\,\mathbf{E}}{c} \\
&= \frac{4\pi\rho}{c}\mathbf{u} - \frac{u\gamma}{c}\frac{\partial \mathbf{E}}{\partial x_1'} \; .
\end{aligned}
\tag{1.130}
$$

On the other hand according to Eq. (1.122) the following holds (recall that $c\beta = u$ and note that $\mathbf{E}$ does not depend on t')

$$
\frac{1}{c}\frac{\partial \mathbf{E}}{\partial t} = -\frac{u\gamma}{c}\frac{\partial \mathbf{E}}{\partial x_1'} \; . \tag{1.131}
$$

The last two results together with Eqs. (1.110) and (1.111) lead to Eq. (1.118). Using the vector identity

$$
\boldsymbol{\nabla} \cdot (\mathbf{A} \times \mathbf{B}) = \mathbf{B} \cdot (\boldsymbol{\nabla} \times \mathbf{A}) - \mathbf{A} \cdot (\boldsymbol{\nabla} \times \mathbf{B}) \; , \tag{1.132}
$$

one finds that [see Eq. (1.109)]

$$
\begin{aligned}
\boldsymbol{\nabla} \cdot \mathbf{B} &= \frac{\boldsymbol{\nabla} \cdot (\mathbf{u} \times \mathbf{E})}{c} \\
&= -\frac{\mathbf{u} \cdot (\boldsymbol{\nabla} \times \mathbf{E})}{c} \\
&= -\frac{u}{c}\left(\frac{\partial E_3}{\partial x_2} - \frac{\partial E_2}{\partial x_3}\right) \; ,
\end{aligned}
\tag{1.133}
$$

or [see Eqs. (1.108), (1.124) and (1.125)]

$$
\boldsymbol{\nabla} \cdot \mathbf{B} = -\frac{u\gamma}{c}\left(\frac{\partial E_3'}{\partial x_2'} - \frac{\partial E_2'}{\partial x_3'}\right) = -\frac{\gamma \mathbf{u} \cdot \left(\boldsymbol{\nabla}' \times \mathbf{E}'\right)}{c} \; . \tag{1.134}
$$

The above result together with Eq. (1.106) lead to Eq. (1.121). Finally, with the help of Eqs. (1.108), (1.109), (1.122), (1.123), (1.124) and (1.125) one obtains

$$
\begin{aligned}
&\boldsymbol{\nabla} \times \mathbf{E} + \frac{1}{c}\frac{\partial \mathbf{B}}{\partial t} \\
&= \left(\frac{\partial E_3}{\partial x_2} - \frac{\partial E_2}{\partial x_3}, \frac{\partial E_1}{\partial x_3} - \frac{\partial E_3}{\partial x_1}, \frac{\partial E_2}{\partial x_1} - \frac{\partial E_1}{\partial x_2}\right) \\
&\quad + \frac{1}{c}\frac{\partial}{\partial t}\frac{\mathbf{u} \times \mathbf{E}}{c} \\
&= \left(\frac{\gamma\partial E_3'}{\partial x_2'} - \frac{\gamma\partial E_2'}{\partial x_3'}, \frac{\partial E_1'}{\partial x_3'} - \frac{\gamma^2\partial E_3'}{\partial x_1'}, \frac{\gamma^2\partial E_2'}{\partial x_1'} - \frac{\partial E_1'}{\partial x_2'}\right) \\
&\quad - \gamma^2\left(\frac{u}{c}\right)^2\left(0, -\frac{\partial E_3'}{\partial x_1'}, \frac{\partial E_2'}{\partial x_1'}\right) \; .
\end{aligned}
\tag{1.135}
$$

By subtracting the term $\gamma\left(\boldsymbol{\nabla}'\times\mathbf{E}'\right)=0$ [see Eq. (1.106)] one obtains

$$
\begin{aligned}
&\boldsymbol{\nabla}\times\mathbf{E}+\frac{1}{c}\frac{\partial\mathbf{B}}{\partial t}-\gamma\left(\boldsymbol{\nabla}'\times\mathbf{E}'\right)\\
&=(1-\gamma)\left(0,\frac{\partial E_1'}{\partial x_3'},-\frac{\partial E_1'}{\partial x_2'}\right)\\
&\left(-\gamma+\gamma^2\left(1-\left(\frac{u}{c}\right)^2\right)\right)\left(0,-\frac{\partial E_3'}{\partial x_1'},\frac{\partial E_2'}{\partial x_1'}\right)\ ,
\end{aligned}
\tag{1.136}
$$

thus [note that $\gamma^2\left(1-\left(\frac{u}{c}\right)^2\right)=1$ and employ again Eq. (1.106)]

$$
\begin{aligned}
&\boldsymbol{\nabla}\times\mathbf{E}+\frac{1}{c}\frac{\partial\mathbf{B}}{\partial t}-\gamma\left(\boldsymbol{\nabla}'\times\mathbf{E}'\right)\\
&=(1-\gamma)\left(0,\frac{\partial E_1'}{\partial x_3'}-\frac{\partial E_3'}{\partial x_1'},\frac{\partial E_2'}{\partial x_1'}-\frac{\partial E_1'}{\partial x_2'}\right)\\
&=0\ ,
\end{aligned}
\tag{1.137}
$$

in agreement with Eq. (1.119).

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 $F=ma$ as the ratio between force F and acceleration a) 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 $c\beta_1\hat{\boldsymbol{\beta}}$ and the relative velocity of S'' with respect to S' is $c\beta_2\hat{\boldsymbol{\beta}}$. Find a Lorentz transformation mapping from S to S''.
2. A uniform and time independent electric field of magnitude E is applied to an electron having charge e and mass m, which is at rest initially at time $t = 0$. Calculate the velocity of the electron $v(t)$ at time $t \geq 0$.
3. A uniform and time independent magnetic field $B\hat{\mathbf{z}}$ is applied to an electron having charge e and mass m. Calculate the period time T of circular motion with velocity v in the xy plane.
4. **The Compton effect** - Consider a photon having energy $E_{\mathrm{p,in}}$ hitting an electron at rest. Calculate the energy of the reflected photon $E_{\mathrm{p,out}}$ for the case where the scattered photon is back-reflected, i.e. its direction is reversed in the process.
5. Show that free electrons can neither emit nor absorb photons.
6. **The Doppler effect** - Consider a plane wave (not necessarily an electromagnetic wave) having the form
$$\psi = \psi_0 \cos\phi \,, \tag{1.138}$$
where the amplitude ψ_0 is a constant, the phase ϕ is given by
$$\phi = \mathbf{k}\cdot\mathbf{r} - \omega t, \tag{1.139}$$
where both wave 3-vector $\mathbf{k} = (k_1, k_2, k_3)$ and angular frequency ω are constants. While the values $\mathbf{k}$ and ω are measured in an inertial frame S, the values $\mathbf{k}'$ and ω' are measured in another inertial frame S' moving at velocity $c\boldsymbol{\beta}$ with respect to S. Calculate $\mathbf{k}'$ and ω'.
7. 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 $v_1 = \mathrm{d}x_1/\mathrm{d}t$, are assumed to vanish at time $t = 0$. Calculate $x_1(t)$.

8. Consider an observer moving along a given trajectory, which in a given inertial frame S is taken to be given by $(ct_{\mathrm{T}}(\tau), x_{\mathrm{T}}(\tau), 0, 0)$, where τ is the proper time, i.e. τ 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

$$t' = \frac{t_1' + t_2'}{2} , \tag{1.140}$$

$$x' = c\frac{t_2' - t_1'}{2} . \tag{1.141}$$

a) Derive relations between the coordinates t and x and the coordinates t' and x' given by Eqs. (1.140) and (1.141).
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 βc 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

$$\beta = \frac{1}{c}\frac{\mathrm{d}x_{\mathrm{T}}}{\mathrm{d}t_{\mathrm{T}}} . \tag{1.142}$$

Show that

$$\frac{t - t'}{t} = \frac{\hat{\beta}\left(x - x_{\mathrm{T}}(0)\right)}{ct} + O\left(\beta^2\right) , \tag{1.143}$$

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

$$\hat{\beta} = \frac{\int_{t-\frac{x-x_{\mathrm{T}}(0)}{c}}^{t+\frac{x-x_{\mathrm{T}}(0)}{c}} \mathrm{d}\tau'\, \beta}{2\frac{x-x_{\mathrm{T}}(0)}{c}} , \tag{1.144}$$

represents the averaged and normalized velocity over the time interval $[t - (x - x_{\mathrm{T}}(0))/c, t + (x - x_{\mathrm{T}}(0))/c]$.

9. **The Unruh-Davies Effect** - Consider an electromagnetic plane wave having amplitude A and wavelength λ. 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.

$$f(\omega') = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \mathrm{d}t' \; f(t') \, e^{i\omega' t'} \; . \tag{1.145}$$

10. **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 $X_n(\tau_n) = (ct_n(\tau_n), \mathbf{r}_n(\tau_n))^{\mathrm{T}}$ be the trajectory in space-time of the n'th point particle, where τ_n is the proper time, i.e. τ_n is the time as being measured by a clock that is carried along with the n'th particle. The current 4-vector $J(X) = (c\rho(X), \mathbf{J}(X))^{\mathrm{T}}$ at space-time point $X = (ct, \mathbf{r})^{\mathrm{T}}$ is defined by

$$J(X) = \sum_n J_n(X) \; , \tag{1.146}$$

where $J_n(X) = (c\rho_n, \mathbf{J}_n)^{\mathrm{T}}$, which represents the contribution of the n'th particle to the total current 4-vector, is taken to be given by

$$J_n(X) = q_n c \int \mathrm{d}\tau_n \frac{\mathrm{d}X_n}{\mathrm{d}\tau_n} \delta(X - X_n(\tau_n)) \; . \tag{1.147}$$

Note that the invariance of the proper time τ_n and the fact that $\mathrm{d}X_n$ is a 4-vector imply that $J(X)$ is a 4-vector. Find expressions for $\rho(X)$ and $\mathbf{J}(X)$.

11. **Dirac equation** - In non-relativistic quantum mechanics, the time evolution of a state vector $|\alpha\rangle$ is governed by the Schrödinger equation

$$i\hbar \frac{\mathrm{d}|\alpha\rangle}{\mathrm{d}t} = \mathcal{H} |\alpha\rangle \; , \tag{1.148}$$

where $\hbar$ is the h-bar Planck's constant and the Hermitian operator $\mathcal{H}$ is the Hamiltonian of the system. Consider a free particle having mass m. For this case the Hamiltonian is given by $\mathcal{H} = \mathbf{p}^2/(2m)$, where $\mathbf{p}$ is the momentum vector operator. In the position representation the Schrödinger equation yields an equation for the wave function $\psi(\mathbf{r}, t)$ given by

$$i\hbar \frac{\mathrm{d}\psi}{\mathrm{d}t} = \frac{(-i\hbar \boldsymbol{\nabla})^2}{2m} \psi \; . \tag{1.149}$$

In view of these relations, one may associate the term $i\hbar(\partial/\partial t)$ with the energy of the particle E (represented by the Hamiltonian $\mathcal{H}$), and the term $-i\hbar\boldsymbol{\nabla} = -i\hbar\left(\partial/\partial x_1, \partial/\partial x_2, \partial/\partial x_3\right)$ with the momentum vector of the particle $\mathbf{p}$. These associations together with the relativistic relation given by Eq. (1.74) suggest the following relation (known as the Klein-Gordon equation)

$$\left(\frac{mc}{\hbar}\right)^2 = -\frac{1}{c^2}\frac{\partial^2}{\partial t^2} + \frac{\partial^2}{\partial x_1^2} + \frac{\partial^2}{\partial x_2^2} + \frac{\partial^2}{\partial x_3^2} \; . \tag{1.150}$$

Consider the following first order equation for ψ

$$\left(i\partial\Gamma - \frac{mc}{\hbar}\right)\psi = 0 \; , \tag{1.151}$$

where the 4-vector Γ is given by $\Gamma = (\gamma_0, \gamma_1, \gamma_2, \gamma_3)^{\mathrm{T}}$, and ∂ is given by Eq. (1.114). This equation was first considered by Dirac as a possible relativistic generalization of the quantum Schrödinger equation. By multiplying Eq. (1.151) by its complex conjugate one obtains

$$\psi^* \left(-i\partial\Gamma - \frac{mc}{\hbar}\right)\left(i\partial\Gamma - \frac{mc}{\hbar}\right)\psi = 0 \; . \tag{1.152}$$

Motivated by the Klein-Gordon relation (1.150), it is required that

$$\begin{aligned} &\left(-i\partial\Gamma - \frac{mc}{\hbar}\right)\left(i\partial\Gamma - \frac{mc}{\hbar}\right) \\ &= \left(\frac{mc}{\hbar}\right)^2 + \frac{1}{c^2}\frac{\partial^2}{\partial t^2} - \frac{\partial^2}{\partial x_1^2} - \frac{\partial^2}{\partial x_2^2} - \frac{\partial^2}{\partial x_3^2} \; . \end{aligned} \tag{1.153}$$

This requirement holds provided that the 4-vector Γ satisfies the following relations (m is treated as a constant)

$$1 = \gamma_0^2 \; , \tag{1.154}$$

$$-1 = \gamma_1^2 = \gamma_2^2 = \gamma_3^2 \; , \tag{1.155}$$

and

$$0 = \gamma_n\gamma_m + \gamma_m\gamma_n \; , \tag{1.156}$$

for $n \neq m$. These relations cannot be all satisfied for the case where γ_0, γ_1, γ_2 and γ_3 are treated as numbers, however, it can be solved when these variables are treated as 4×4 matrices. Find 4×4 matrix representations for γ_0, γ_1, γ_2 and γ_3, and use these representations to derive the Dirac equation for the 4-vector wavefunction $(\psi_0, \psi_1, \psi_2, \psi_3)$.

1.7 Solutions

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

$$\Lambda = \exp\left(-\left(\kappa_1+\kappa_2\right)\hat{\boldsymbol{\beta}}\cdot\Sigma\right) , \tag{1.157}$$

where

$$\kappa_{1,2} = \tanh^{-1}\beta_{1,2} . \tag{1.158}$$

Using the identity

$$\tanh\left(\kappa_1+\kappa_2\right) = \frac{\tanh\left(\kappa_1\right)+\tanh\left(\kappa_2\right)}{1+\tanh\left(\kappa_1\right)\tanh\left(\kappa_2\right)} , \tag{1.159}$$

one finds that

$$\beta = \frac{\beta_1+\beta_2}{1+\beta_1\beta_2} , \tag{1.160}$$

where

$$\beta = \tanh\left(\kappa_1+\kappa_2\right) . \tag{1.161}$$

2. The electron momentum p is related to its velocity v by $p = m\gamma v$ [see Eq. (1.70)]. The solution of $eE = \mathrm{d}p/\mathrm{d}t$ [see Eq. (1.79)] for the given initial condition $p\left(t=0\right) = 0$ is given by $p = eEt$, and thus [see Eq. (1.69)]

$$\frac{mv}{\sqrt{1-\frac{v^2}{c^2}}} = eEt , \tag{1.162}$$

and therefore

$$v = \frac{\frac{eEt}{mc}}{\sqrt{1+\left(\frac{eEt}{mc}\right)^2}}c . \tag{1.163}$$

3. The equation of motion (1.79) for this case reads [see Eq. (1.107)]

$$e\frac{\mathbf{v}\times\mathbf{B}}{c} = \frac{\mathrm{d}\mathbf{p}}{\mathrm{d}t} , \tag{1.164}$$

where $\mathbf{p} = m\gamma\mathbf{v}$ [see Eq. (1.70)]. For circular motion in the xy plane with velocity v and period time $T = 2\pi/\omega$ (where ω is the angular frequency)

$$\frac{evB}{c} = m\gamma\omega v , \tag{1.165}$$

hence [see Eq. (1.69)]

$$\omega = \frac{2\pi}{T} = \frac{eB}{cm}\sqrt{1-\frac{v^2}{c^2}} . \tag{1.166}$$

4. Energy-momentum conservation implies [see Eqs. (1.74), (1.75) and (1.76)]

$$E_{\mathrm{p,in}} + m_{\mathrm{e}}c^2 = E_{\mathrm{p,out}} + \sqrt{m_{\mathrm{e}}^2 c^4 + p^2 c^2}\ , \tag{1.167}$$

$$\frac{E_{\mathrm{p,in}}}{c} = -\frac{E_{\mathrm{p,out}}}{c} + p\ , \tag{1.168}$$

where m_{e} is the electron mass and p is the momentum of the scattered electron. By solving for the unknowns $E_{\mathrm{p,out}}$ and p one obtains

$$\frac{E_{\mathrm{p,out}}}{m_{\mathrm{e}}c^2} = \frac{\frac{E_{\mathrm{p,in}}}{m_{\mathrm{e}}c^2}}{1 + 2\frac{E_{\mathrm{p,in}}}{m_{\mathrm{e}}c^2}}\ . \tag{1.169}$$

5. 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. (1.74), (1.75) and (1.76)]

$$\mathbf{p}_{\mathrm{p}} = \mathbf{p}_{\mathrm{e}}\ , \tag{1.170}$$

$$p_{\mathrm{p}}c + m_{\mathrm{e}}c^2 = \sqrt{m_{\mathrm{e}}^2 c^4 + p_{\mathrm{e}}^2 c^2}\ , \tag{1.171}$$

and for the case of photon emission that

$$0 = \mathbf{p}_{\mathrm{p}} + \mathbf{p}_{\mathrm{e}}\ , \tag{1.172}$$

$$m_{\mathrm{e}}c^2 = p_{\mathrm{p}}c + \sqrt{m_{\mathrm{e}}^2 c^4 + p_{\mathrm{e}}^2 c^2}\ , \tag{1.173}$$

where $\mathbf{p}_{\mathrm{p}}$ and $\mathbf{p}_{\mathrm{e}}$ 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 $p_{\mathrm{p}} = p_{\mathrm{e}} = 0$.

6. Consider the 4-vector K, which is defined by

$$K = -\partial\eta\phi = \left(\frac{\omega}{c}, k_1, k_2, k_3\right)\ , \tag{1.174}$$

where $\partial = \left(c^{-1}\partial/\partial t, \partial/\partial x_1, \partial/\partial x_2, \partial/\partial x_3\right)$ [see Eq. (1.114)] and where $\phi = \mathbf{k}\cdot\mathbf{r} - \omega t$ [see Eq. (1.139)]. Since ∂ is transformed according to $\partial' = \partial\Lambda^{-1}$ [see Eq. (1.115)] and because ϕ is expected to be Lorentz invariant (explain why) one concludes that the 4-vector K is transformed according to [see Eq. (1.21)]

$$K' = -\partial'\eta\phi = -\partial\Lambda^{-1}\eta\phi = -\partial\eta\Lambda^{\mathrm{T}}\phi\ . \tag{1.175}$$

The coordinates of frame S are chosen such that the velocity $c\boldsymbol{\beta}$ is pointing in the x_1 direction. For this case Eq. (1.175) becomes [see Eqs. (1.14) and (1.26)]

$$\begin{aligned} K' &= \left(\gamma\left(-\frac{1}{c}\frac{\partial\phi}{\partial t} - \beta\frac{\partial\phi}{\partial x_1}\right), \gamma\left(\frac{\beta}{c}\frac{\partial\phi}{\partial t} + \frac{\partial\phi}{\partial x_1}\right), \frac{\partial\phi}{\partial x_2}, \frac{\partial\phi}{\partial x_3}\right) \\ &= \left(\gamma\left(\frac{\omega}{c} - \beta k_1\right), \gamma\left(-\frac{\beta\omega}{c} + k_1\right), k_2, k_3\right)\ , \end{aligned} \tag{1.176}$$

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

$$\omega' = \gamma\omega\left(1 - \frac{c\beta k_1}{\omega}\right) , \tag{1.177}$$

and

$$\mathbf{k}' = \left(\gamma\left(k_1 - \frac{\beta\omega}{c}\right), k_2, k_3\right) . \tag{1.178}$$

Let θ (θ') be the angle between $\mathbf{k}$ ($\mathbf{k}'$) and $\boldsymbol{\beta}$

$$\theta = \cos^{-1}\frac{k_1}{k} , \tag{1.179}$$

$$\theta' = \cos^{-1}\frac{k_1'}{k'} , \tag{1.180}$$

where

$$k = \sqrt{k_1^2 + k_2^2 + k_3^2} , \tag{1.181}$$

$$k' = \sqrt{k_1'^2 + k_2'^2 + k_3'^2} . \tag{1.182}$$

Using this notation Eq. (1.177) can be rewritten as

$$\frac{\omega'}{\omega} = \frac{1 - \frac{ck}{\omega}\beta\cos\theta}{\sqrt{1-\beta^2}} . \tag{1.183}$$

The angle θ' can be found from Eq. (1.178)

$$\tan\theta' = \frac{\sin\theta\sqrt{1-\beta^2}}{\cos\theta - \frac{\beta\omega}{ck}} . \tag{1.184}$$

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

7. In the instantaneous rest frame of the particle the velocity 4-vector of the particle U is given by [see Eq. (1.64)]

$$U = \frac{\mathrm{d}X}{\mathrm{d}\tau} = (c, 0, 0, 0) . \tag{1.185}$$

and the acceleration 4-vector A is given by

$$A = \frac{\mathrm{d}^2X}{\mathrm{d}\tau^2} = (0, a, 0, 0) . \tag{1.186}$$

Using the Lorentz transformation for U and A one obtains

$$\frac{\mathrm{d}}{\mathrm{d}\tau}\begin{pmatrix} ct \\ x_1 \end{pmatrix} = B_1\begin{pmatrix} c \\ 0 \end{pmatrix} , \tag{1.187}$$

and

$$\frac{\mathrm{d}^2}{\mathrm{d}\tau^2}\begin{pmatrix} ct \\ x_1 \end{pmatrix} = B_1 \begin{pmatrix} 0 \\ a \end{pmatrix} , \tag{1.188}$$

where [see Eq. (1.26)]

$$B_1 = \gamma \begin{pmatrix} 1 & -\beta \\ -\beta & 1 \end{pmatrix} , \tag{1.189}$$

$\beta = -v_1/c$ and $\gamma = 1/\sqrt{1 - v_1^2/c^2}$. With the help of Eq. (1.189) Eq. (1.187) becomes

$$\begin{pmatrix} \frac{\mathrm{d}t}{\mathrm{d}\tau} \\ \frac{\mathrm{d}x_1}{\mathrm{d}t} \end{pmatrix} = \begin{pmatrix} \gamma \\ v_1 \end{pmatrix} . \tag{1.190}$$

Substituting Eq. (1.187) into Eq. (1.188) yields

$$\begin{pmatrix} \frac{\mathrm{d}\gamma}{\mathrm{d}t} \\ \frac{\mathrm{d}(\gamma v_1)}{\mathrm{d}t} \end{pmatrix} = \begin{pmatrix} \frac{v_1 a}{c^2} \\ a \end{pmatrix} . \tag{1.191}$$

The second equation of (1.191), which reads

$$\frac{\mathrm{d}\left(\frac{v_1}{\sqrt{1 - v_1^2/c^2}}\right)}{\mathrm{d}t} = a , \tag{1.192}$$

can be solved using the transformation

$$\frac{v_1}{c} = \tanh s , \tag{1.193}$$

which yields (recall that $1 - \tanh^2 s = 1/\cosh^2 s$)

$$\frac{\mathrm{d}\left(\sinh s\right)}{\mathrm{d}t} = \frac{a}{c} . \tag{1.194}$$

Integration and employing the initial conditions at time $t = 0$ lead to [see Eq. (1.193)]

$$\frac{v_1}{\sqrt{1 - \frac{v_1^2}{c^2}}} = at , \tag{1.195}$$

and thus

$$v_1\left(t\right) = \frac{at}{\sqrt{1 + \left(\frac{at}{c}\right)^2}} . \tag{1.196}$$

By integrating Eq. (1.196) one obtains

$$x_1(t) = \frac{c^2}{a}\left(\sqrt{1+\frac{a^2t^2}{c^2}} - 1\right) . \tag{1.197}$$

Alternatively, the position x_1 and velocity v_1 can be expressed as a function of the proper time τ as follows. With the help of Eq. (1.196) the first equation of (1.190) becomes

$$\frac{\mathrm{d}t}{\mathrm{d}\tau} = \sqrt{1+\left(\frac{at}{c}\right)^2} , \tag{1.198}$$

and thus

$$t = \frac{c}{a}\sinh\frac{a\tau}{c} . \tag{1.199}$$

The last result (1.199) together with Eqs. (1.196) and (1.197) yield

$$v_1(\tau) = c\tanh\frac{a\tau}{c} , \tag{1.200}$$

and

$$x_1(\tau) = \frac{c^2}{a}\left(\cosh\frac{a\tau}{c} - 1\right) . \tag{1.201}$$

8. The light signal sent by the observer connects the space-time points $(ct_{\mathrm{T}}(t_1'), x_{\mathrm{T}}(t_1'), 0, 0)$ and $(ct, x, 0, 0)$, and the back reflected light signal connects the space-time points $(ct, x, 0, 0)$ and $(ct_{\mathrm{T}}(t_2'), x_{\mathrm{T}}(t_2'), 0, 0)$, and thus

$$\frac{x - x_{\mathrm{T}}(t_1')}{t - t_{\mathrm{T}}(t_1')} = c , \tag{1.202}$$

$$\frac{x_{\mathrm{T}}(t_2') - x}{t_{\mathrm{T}}(t_2') - t} = -c , \tag{1.203}$$

a) With the help of Eqs. (1.140) and (1.141) the relations (1.202) and (1.203) can be rewritten as

$$\frac{x - x_{\mathrm{T}}\left(t' - \frac{x'}{c}\right)}{t - t_{\mathrm{T}}\left(t' - \frac{x'}{c}\right)} = c , \tag{1.204}$$

$$\frac{x_{\mathrm{T}}\left(t' + \frac{x'}{c}\right) - x}{t_{\mathrm{T}}\left(t' + \frac{x'}{c}\right) - t} = -c , \tag{1.205}$$

or

$$x - ct = x_{\mathrm{T}}\left(t' - \frac{x'}{c}\right) - ct_{\mathrm{T}}\left(t' - \frac{x'}{c}\right) , \tag{1.206}$$

$$x + ct = x_{\mathrm{T}}\left(t' + \frac{x'}{c}\right) + ct_{\mathrm{T}}\left(t' + \frac{x'}{c}\right) . \tag{1.207}$$

b) For this case $t_{\mathrm{T}}(\tau) = \tau$ and $x_{\mathrm{T}}(\tau) = 0$, and thus Eqs. (1.206) and (1.207) become
$$x - ct = -c\left(t' - \frac{x'}{c}\right) , \tag{1.208}$$
$$x + ct = c\left(t' + \frac{x'}{c}\right) . \tag{1.209}$$
The solution is $t' = t$ and $x' = x$.

c) For this case $t_{\mathrm{T}}(\tau) = \gamma\tau$ and $x_{\mathrm{T}}(\tau) = \beta ct = \beta c\gamma\tau$ [see Eq. (1.34)], where
$$\gamma = \frac{1}{\sqrt{1-\beta^2}} , \tag{1.210}$$
and thus Eqs. (1.206) and (1.207) become
$$x - ct = \beta c\gamma\left(t' - \frac{x'}{c}\right) - c\gamma\left(t' - \frac{x'}{c}\right) , \tag{1.211}$$
$$x + ct = \beta c\gamma\left(t' + \frac{x'}{c}\right) + c\gamma\left(t' + \frac{x'}{c}\right) . \tag{1.212}$$
The solution can be written as [compare with Eq. (1.25)]
$$\begin{pmatrix} ct' \\ x' \end{pmatrix} = \gamma \begin{pmatrix} 1 & -\beta \\ -\beta & 1 \end{pmatrix} \begin{pmatrix} ct \\ x \end{pmatrix} . \tag{1.213}$$

d) For this case [see Eqs. (1.199) and (1.201)]
$$t_{\mathrm{T}}(\tau) = \frac{c}{a}\sinh\frac{a\tau}{c} , \tag{1.214}$$
$$x_{\mathrm{T}}(\tau) = \frac{c^2}{a}\left(\cosh\frac{a\tau}{c} - 1\right) , \tag{1.215}$$
and thus Eqs. (1.206) and (1.207) become
$$x - ct = \frac{c^2}{a}\left(e^{\frac{a\left(\frac{x'}{c} - t'\right)}{c}} - 1\right) , \tag{1.216}$$
$$x + ct = \frac{c^2}{a}\left(e^{\frac{a\left(\frac{x'}{c} + t'\right)}{c}} - 1\right) . \tag{1.217}$$
The solution is given by
$$x = \frac{c^2}{a}\left[\exp\left(\frac{ax'}{c^2}\right)\cosh\left(\frac{at'}{c}\right) - 1\right] , \tag{1.218}$$
$$ct = \frac{c^2}{a}\exp\left(\frac{ax'}{c^2}\right)\sinh\left(\frac{at'}{c}\right) . \tag{1.219}$$
To lowest nonvanishing order in a
$$x = \left(1 + \frac{x'}{2\frac{c^2}{a}}\right)x' + \frac{at'^2}{2} + O\left(a^2\right) , \tag{1.220}$$
$$ct = ct'\left(1 + \frac{x'}{\frac{c^2}{a}}\right) + O\left(a^2\right) . \tag{1.221}$$

The inverse transformation is given by

$$\begin{aligned} x' &= \frac{c^2}{2a} \log \left[\left(1 + \frac{x}{\frac{c^2}{a}} \right)^2 - \left(\frac{at}{c} \right)^2 \right] \\ &= \left(1 - \frac{x}{2\frac{c^2}{a}} \right) x - \frac{at^2}{2} + O\left(a^2\right) \ , \end{aligned} \tag{1.222}$$

$$\begin{aligned} ct' &= \frac{c^2}{a} \tanh^{-1} \frac{ct}{x + \frac{c^2}{a}} \\ &= ct \left(1 - \frac{x}{\frac{c^2}{a}} \right) + O\left(a^2\right) \ . \end{aligned} \tag{1.223}$$

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

$$\frac{\mathrm{d}}{\mathrm{d}\tau} \begin{pmatrix} ct_{\mathrm{T}} \\ x_{\mathrm{T}} \end{pmatrix} = \Lambda(-\beta) \begin{pmatrix} c \\ 0 \end{pmatrix} \ , \tag{1.224}$$

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

$$\Lambda(-\beta) = \gamma \begin{pmatrix} 1 & \beta \\ \beta & 1 \end{pmatrix} \ , \tag{1.225}$$

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

$$v = \frac{\mathrm{d}x_{\mathrm{T}}}{\mathrm{d}t_{\mathrm{T}}} \ . \tag{1.226}$$

Integrating Eq. (1.224), which can be rewritten as

$$\frac{\mathrm{d}t_{\mathrm{T}}}{\mathrm{d}\tau} = \gamma \ , \tag{1.227}$$

$$\frac{\mathrm{d}x_{\mathrm{T}}}{\mathrm{d}\tau} = c\gamma\beta \ , \tag{1.228}$$

with the assumed initial condition $t_{\mathrm{T}}(\tau = 0) = 0$, yields

$$t_{\mathrm{T}}(\tau) = \int_0^\tau \mathrm{d}\tau' \, \gamma(\tau') \ , \tag{1.229}$$

$$x_{\mathrm{T}}(\tau) = x_{\mathrm{T}}(0) + c \int_0^\tau \mathrm{d}\tau' \, \gamma(\tau') \, \beta(\tau') \ . \tag{1.230}$$

With the help of Eqs. (1.229) and (1.230) one finds that

$$x_{\mathrm{T}}(\tau) \pm ct_{\mathrm{T}}(\tau) = x_{\mathrm{T}}(0) + c\int_0^{\tau} \mathrm{d}\tau' \, \gamma(\beta \pm 1) \ , \tag{1.231}$$

and thus Eqs. (1.206) and (1.207) can be rewritten as

$$\frac{x - x_{\mathrm{T}}(0)}{c} - t = \int_0^{t' - \frac{x'}{c}} \mathrm{d}\tau' \, \gamma(\beta - 1) \ , \tag{1.232}$$

$$\frac{x - x_{\mathrm{T}}(0)}{c} + t = \int_0^{t' + \frac{x'}{c}} \mathrm{d}\tau' \, \gamma(\beta + 1) \ , \tag{1.233}$$

or alternatively, with the help of the relations

$$-\gamma + \beta\gamma = -\sqrt{\frac{1-\beta}{1+\beta}} \ , \tag{1.234}$$

$$-\gamma - \beta\gamma = -\sqrt{\frac{1+\beta}{1-\beta}} \ , \tag{1.235}$$

as

$$t' - \frac{x'}{c} = t - \frac{x - x_{\mathrm{T}}(0)}{c} + \int_0^{t' - \frac{x'}{c}} \mathrm{d}\tau' \left(1 - \sqrt{\frac{1-\beta}{1+\beta}}\right) \ , \tag{1.236}$$

$$t' + \frac{x'}{c} = t + \frac{x - x_{\mathrm{T}}(0)}{c} + \int_0^{t' + \frac{x'}{c}} \mathrm{d}\tau' \left(1 - \sqrt{\frac{1+\beta}{1-\beta}}\right) \ . \tag{1.237}$$

In the limit of a stationary observer having a vanishing velocity both integrals on the right hand sides of Eqs. (1.236) and (1.237) vanish. The solutions of Eqs. (1.236) and (1.237) in this limit provides approximated values for the upper limits of these integrals, which can be used to turn Eqs. (1.236) and (1.237) into

$$\begin{aligned} t' - \frac{x'}{c} &= t - \frac{x - x_{\mathrm{T}}(0)}{c} \\ &\quad + \int_0^{t - \frac{x - x_{\mathrm{T}}(0)}{c}} \mathrm{d}\tau' \left(1 - \sqrt{\frac{1-\beta}{1+\beta}}\right) + O\left(\beta^2\right) \ , \end{aligned} \tag{1.238}$$

and

$$\begin{aligned} t' + \frac{x'}{c} &= t + \frac{x - x_{\mathrm{T}}(0)}{c} \\ &\quad + \int_0^{t + \frac{x - x_{\mathrm{T}}(0)}{c}} \mathrm{d}\tau' \left(1 - \sqrt{\frac{1+\beta}{1-\beta}}\right) + O\left(\beta^2\right) \ . \end{aligned} \tag{1.239}$$

By adding the above equations one obtains

$$\begin{aligned}\frac{t-t'}{t} &= -1+\frac{1}{2t}\int_0^{t-\frac{x-x_{\mathrm{T}}(0)}{c}} \mathrm{d}\tau' \sqrt{\frac{1-\beta}{1+\beta}} \\ &\quad +\frac{1}{2t}\int_0^{t+\frac{x-x_{\mathrm{T}}(0)}{c}} \mathrm{d}\tau' \sqrt{\frac{1+\beta}{1-\beta}} .\end{aligned} \tag{1.240}$$

The expansions

$$\sqrt{\frac{1-\beta}{1+\beta}} = 1-\beta+O\left(\beta^2\right) , \tag{1.241}$$

$$\sqrt{\frac{1+\beta}{1-\beta}} = 1+\beta+O\left(\beta^2\right) , \tag{1.242}$$

lead to Eq. (1.143). Note that for the case of a constant β Eq. (1.143) yields

$$\frac{t-t'}{t} = \frac{\beta\left(x-x_{\mathrm{T}}(0)\right)}{ct} , \tag{1.243}$$

whereas the exact result (1.25) is

$$\frac{t-t'}{t} = \frac{\gamma\beta\left(x-x_{\mathrm{T}}(0)\right)}{ct}+1-\gamma , \tag{1.244}$$

in agreement with the approximated result only to first order in β.

9. The plane wave can be expressed using either the inertial frame coordinates t and x_1

$$f\left(t,x_1\right) = Ae^{i\omega_0\left(\frac{x_1}{c}-t\right)} , \tag{1.245}$$

where

$$\omega_0 = \frac{2\pi c}{\lambda} , \tag{1.246}$$

or the observer's coordinates t' and x_1' [see Eq. (1.216)]

$$\begin{aligned}f\left(t',x_1'\right) &= A\exp\left(\frac{i\omega_0 c}{a}\left(e^{\frac{a}{c}\left(\frac{x_1'}{c}-t'\right)}-1\right)\right) \\ &= Ae^{-\frac{i\omega_0}{\omega_{\mathrm{a}}}}e^{i\alpha e^{-\omega_{\mathrm{a}}t'}} ,\end{aligned} \tag{1.247}$$

where

$$\omega_{\mathrm{a}} = \frac{a}{c}^{i\alpha e^{-\omega_{\mathrm{a}}t'}} ,$$

and where

$$\alpha = \frac{\omega_0}{\omega_{\mathrm{a}}} e^{\frac{\omega_{\mathrm{a}} x_1'}{c}} \; . \tag{1.248}$$

By employing the integral variable transformation $z = \alpha e^{-\omega_{\mathrm{a}} t'}$ one obtains [see Eq. (1.145)]

$$\begin{aligned} f(\omega') &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \mathrm{d}t' \; f(t', x_1')\, e^{i\omega' t'} \\ &= \frac{A e^{-\frac{i\omega_0}{\omega_{\mathrm{a}}}} \alpha^{\frac{i\omega'}{\omega_{\mathrm{a}}}}}{\sqrt{2\pi}\omega_{\mathrm{a}}} \int_0^{\infty} \mathrm{d}z \; e^{iz} z^{-\frac{i\omega'}{\omega_{\mathrm{a}}}-1} \; , \end{aligned} \tag{1.249}$$

and thus

$$|f(\omega')|^2 = \frac{|A|^2}{\omega_{\mathrm{a}}\omega'} n_{\mathrm{BE}} \left(\frac{2\pi\omega'}{\omega_{\mathrm{a}}} \right) \; , \tag{1.250}$$

where

$$n_{\mathrm{BE}}(\epsilon) = \frac{1}{e^{\epsilon} - 1} \tag{1.251}$$

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

$$T_{\mathrm{UD}} = \frac{\hbar a}{2\pi k_{\mathrm{B}} c} \; , \tag{1.252}$$

where k_{B} is the Boltzmann's constant, the result can be rewritten as

$$|f(\omega')|^2 = \frac{|A|^2}{\omega_{\mathrm{a}}\omega'} n_{\mathrm{BE}} \left(\frac{\hbar\omega'}{k_{\mathrm{B}} T_{\mathrm{UD}}} \right) \; . \tag{1.253}$$

10. With the help of Eqs. (1.146) and (1.147) one obtains

$$\rho(t, \mathbf{r}) = \sum_n q_n \int \mathrm{d}\tau_n \frac{\mathrm{d}t_n}{\mathrm{d}\tau_n} \delta(t - t_n(\tau_n))\, \delta(\mathbf{r} - \mathbf{r}_n(\tau_n)) \; , \tag{1.254}$$

and

$$\mathbf{J}(t, \mathbf{r}) = \sum_n q_n \int \mathrm{d}\tau_n \frac{\mathrm{d}\mathbf{r}_n}{\mathrm{d}\tau_n} \delta(t - t_n(\tau_n))\, \delta(\mathbf{r} - \mathbf{r}_n(\tau_n)) \; , \tag{1.255}$$

thus with the help of the relation

$$\delta(t - t_n(\tau_n)) = \left(\left| \frac{\mathrm{d}t_n}{\mathrm{d}\tau_n} \right| \right)^{-1} \delta(\tau_n - \tau_n(t_n)) \; , \tag{1.256}$$

one finds that

$$\rho(t,\mathbf{r}) = \sum_n q_n \delta(\mathbf{r} - \mathbf{r}_n(t)) \ , \tag{1.257}$$

and

$$\mathbf{J}(t,\mathbf{r}) = \sum_n q_n \mathbf{v}_n(t)\, \delta(\mathbf{r} - \mathbf{r}_n(t)) \ , \tag{1.258}$$

where $\mathbf{r}_n(t)$ and $\mathbf{v}_n(t)$ are the location and velocity, respectively, of the n'th particle at time t.

11. The Pauli matrices σ_1, σ_2 and σ_3, which are given by

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} ,\ \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} ,\ \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \ , \tag{1.259}$$

satisfy the following relations

$$\sigma_1^2 = \sigma_2^2 = \sigma_3^2 = \hat{1} \ , \tag{1.260}$$

where

$$\hat{1} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \ , \tag{1.261}$$

and

$$\sigma_n \sigma_m + \sigma_m \sigma_n = 0 \ , \tag{1.262}$$

for $n \neq m$, and thus, in a block form, γ_0, γ_1, γ_2 and γ_3 can be taken to be given by

$$\gamma_0 = \left(\begin{array}{|c|c|} \hline \hat{1} & 0 \\ \hline 0 & -\hat{1} \\ \hline \end{array}\right) \ , \tag{1.263}$$

and

$$\gamma_n = \left(\begin{array}{|c|c|} \hline 0 & \sigma_n \\ \hline -\sigma_n & 0 \\ \hline \end{array}\right) \ , \tag{1.264}$$

for $n \in \{1,2,3\}$. With these 4×4 matrix representations the Dirac equation (1.151) becomes

$$\begin{pmatrix} \frac{i}{c}\frac{\partial}{\partial t} - \frac{mc}{\hbar} & 0 & i\frac{\partial}{\partial x_3} & i\frac{\partial}{\partial x_1} + \frac{\partial}{\partial x_2} \\ 0 & \frac{i}{c}\frac{\partial}{\partial t} - \frac{mc}{\hbar} & i\frac{\partial}{\partial x_1} - \frac{\partial}{\partial x_2} & -i\frac{\partial}{\partial x_3} \\ -i\frac{\partial}{\partial x_3} & -i\frac{\partial}{\partial x_1} - \frac{\partial}{\partial x_2} & -\frac{i}{c}\frac{\partial}{\partial t} - \frac{mc}{\hbar} & 0 \\ -i\frac{\partial}{\partial x_1} + \frac{\partial}{\partial x_2} & i\frac{\partial}{\partial x_3} & 0 & -\frac{i}{c}\frac{\partial}{\partial t} - \frac{mc}{\hbar} \end{pmatrix} \begin{pmatrix} \psi_0 \\ \psi_1 \\ \psi_2 \\ \psi_3 \end{pmatrix} = 0 \ . \tag{1.265}$$

2. Maxwell's Equations in Matter

In this chapter the macroscopic Maxwell's equations, which are used to describe 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 ρ and $\mathbf{J}$ appearing in the Maxwell's equations (1.118), (1.119), (1.120) and (1.121) by their spacial average according to the following procedure

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

where ΔV is the averaging volume, which is chosen such that $d_{\mathrm{A}} \ll \Delta V^{1/3} \ll \lambda$, where d_{A} is the characteristic distance between atoms in the matter and where λ is the characteristic wavelength of electromagnetic fields. The averaged fields and sources satisfy the same set of Maxwell's equations (1.118), (1.119), (1.120) and (1.121)

$$\boldsymbol{\nabla}\times\langle\mathbf{B}\rangle = \frac{4\pi}{c}\langle\mathbf{J}_{\mathrm{total}}\rangle + \frac{1}{c}\frac{\partial\langle\mathbf{E}\rangle}{\partial t} \ , \tag{2.2}$$

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

$$\boldsymbol{\nabla}\cdot\langle\mathbf{E}\rangle = 4\pi\langle\rho_{\mathrm{total}}\rangle \ , \tag{2.4}$$

$$\boldsymbol{\nabla}\cdot\langle\mathbf{B}\rangle = 0 \ . \tag{2.5}$$

Note that the label 'total' has been added as a subscript to ρ and $\mathbf{J}$. This is done because in what follows the charge density $\rho = \rho_{\mathrm{total}}$ and current density $\mathbf{J} = \mathbf{J}_{\mathrm{total}}$ 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 ρ_{total} is decomposed as

$$\rho_{\mathrm{total}} = \rho_{\mathrm{ext}} + \rho_{\mathrm{pol}} \,, \tag{2.6}$$

where ρ_{pol}, which represents the contribution due to dielectric polarization, is given by

$$\rho_{\mathrm{pol}} = -\boldsymbol{\nabla} \cdot \mathbf{P} \,, \tag{2.7}$$

and ρ_{ext} represents all other contributions. The total current density $\mathbf{J}_{\mathrm{total}}$ is decomposed as

$$\mathbf{J}_{\mathrm{total}} = \mathbf{J}_{\mathrm{cond}} + \mathbf{J}_{\mathrm{bound}} + \mathbf{J}_{\mathrm{ext}} \,, \tag{2.8}$$

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

$$\mathbf{J}_{\mathrm{bound}} = \mathbf{J}_{\mathrm{pol}} + \mathbf{J}_{\mathrm{mag}} \,, \tag{2.9}$$

the term $\mathbf{J}_{\mathrm{pol}}$, which is given by

$$\mathbf{J}_{\mathrm{pol}} = \frac{\partial \mathbf{P}}{\partial t} \,, \tag{2.10}$$

represents the contribution of dielectric polarization, the term $\mathbf{J}_{\mathrm{mag}}$, which is given by

$$\mathbf{J}_{\mathrm{mag}} = c\boldsymbol{\nabla} \times \mathbf{M} \,, \tag{2.11}$$

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

$$\boldsymbol{\nabla} \times \mathbf{H} = \frac{4\pi}{c} \left(\mathbf{J}_{\mathrm{ext}} + \mathbf{J}_{\mathrm{cond}} \right) + \frac{1}{c} \frac{\partial \mathbf{D}}{\partial t} \,, \tag{2.12}$$

$$\boldsymbol{\nabla} \times \mathbf{E} = -\frac{1}{c} \frac{\partial \mathbf{B}}{\partial t} \,, \tag{2.13}$$

$$\boldsymbol{\nabla} \cdot \mathbf{D} = 4\pi \rho_{\mathrm{ext}} \,, \tag{2.14}$$

$$\boldsymbol{\nabla} \cdot \mathbf{B} = 0 \,. \tag{2.15}$$

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

$$\boldsymbol{\nabla} \cdot \mathbf{E} = 4\pi \rho_{\mathrm{total}} \,, \tag{2.16}$$

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

$$\mathbf{D} = \mathbf{E} + 4\pi \mathbf{P} \,, \tag{2.17}$$

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

$$\mathbf{H} = \mathbf{B} - 4\pi \mathbf{M} \,, \tag{2.18}$$

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

$$\nabla \times \mathbf{H} = \frac{4\pi}{c}\left(\mathbf{J}_{\text{total}} - \mathbf{J}_{\text{mag}}\right) + \frac{1}{c}\frac{\partial \mathbf{E}}{\partial t} , \tag{2.19}$$

and

$$\nabla \times \mathbf{B} = \frac{4\pi}{c}\mathbf{J}_{\text{total}} + \frac{1}{c}\frac{\partial \mathbf{E}}{\partial t} . \tag{2.20}$$

Exercise 2.1.1. current conservation - Show that

$$\nabla \cdot \left(\mathbf{J}_{\text{ext}} + \mathbf{J}_{\text{cond}}\right) + \frac{\partial \rho_{\text{ext}}}{\partial t} = 0 . \tag{2.21}$$

Solution 2.1.1. This relation, which is the in-matter version of the continuity equation (1.117), can be proven by applying ∇ 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. $\nabla \times \mathbf{E} = (-1/c)\,\partial \mathbf{B}/\partial t$ (2.13) and $\nabla \cdot \mathbf{B} = 0$ (2.15) , which are identical to the Maxwell's equations in vacuum (1.119) and (1.121), 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 $\nabla \times \mathbf{B} = (4\pi/c)\,\mathbf{J} + (1/c)\,\partial \mathbf{E}/\partial t$ (1.118) and $\nabla \cdot \mathbf{E} = 4\pi\rho$ (1.120) by the transformation $\mathbf{E} \rightarrow \mathbf{D}$, $\mathbf{B} \rightarrow \mathbf{H}$ and $J \rightarrow J_{\text{ext}}$, where J_{ext} is defined by [compare with Eq. (1.112)]

$$J_{\text{ext}} = \left(c\rho_{\text{ext}}, \mathbf{J}_{\text{ext}} + \mathbf{J}_{\text{cond}}\right)^{\mathrm{T}} . \tag{2.22}$$

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

$$\mathbf{B} = \nabla \times \mathbf{A} . \tag{2.23}$$

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

$$\nabla \times \left(\mathbf{E} + \frac{1}{c}\frac{\partial \mathbf{A}}{\partial t}\right) = 0 , \tag{2.24}$$

which implies the existence of a scalar ϕ such that

$$\mathbf{E} = -\nabla\phi - \frac{1}{c}\frac{\partial \mathbf{A}}{\partial t} . \tag{2.25}$$

When the fields $\mathbf{E}$ and $\mathbf{B}$ are expressed in terms of ϕ and $\mathbf{A}$ the Maxwell's equations (2.13) and (2.15) are guarantee to be satisfied provided that ϕ 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 A is defined by

$$A = (\phi, A_1, A_2, A_3)^{\mathrm{T}} = (\phi, \mathbf{A})^{\mathrm{T}} \ . \tag{2.26}$$

The quantity ∂A, which is given by

$$\partial A = \frac{1}{c}\frac{\partial \phi}{\partial t} + \boldsymbol{\nabla} \cdot \mathbf{A} \ , \tag{2.27}$$

is Lorentz invariant provided that A is Lorentz transformed according to $A' = \Lambda A$ [see Eq. (1.115)].

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

$$\partial^{\mathrm{T}} A^{\mathrm{T}} \eta - \left(\partial^{\mathrm{T}} A^{\mathrm{T}} \eta\right)^{\mathrm{T}} = \hat{F} \ , \tag{2.28}$$

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

$$\hat{F} = \begin{pmatrix} 0 & E_1 & E_2 & E_3 \\ -E_1 & 0 & -B_3 & B_2 \\ -E_2 & B_3 & 0 & -B_1 \\ -E_3 & -B_2 & B_1 & 0 \end{pmatrix} \ . \tag{2.29}$$

Proof. The following holds [see Eqs. (1.14), (1.114) and (2.26)]

$$\partial^{\mathrm{T}} A^{\mathrm{T}} \eta = \begin{pmatrix} \frac{1}{c}\frac{\partial}{\partial t} \\ \frac{\partial}{\partial x_1} \\ \frac{\partial}{\partial x_2} \\ \frac{\partial}{\partial x_3} \end{pmatrix} \left(\phi \ , -A_1 \ , -A_2 \ , -A_3 \right) = \begin{pmatrix} \frac{1}{c}\frac{\partial \phi}{\partial t} & -\frac{1}{c}\frac{\partial A_1}{\partial t} & -\frac{1}{c}\frac{\partial A_2}{\partial t} & -\frac{1}{c}\frac{\partial A_3}{\partial t} \\ \frac{\partial \phi}{\partial x_1} & -\frac{\partial A_1}{\partial x_1} & -\frac{\partial A_2}{\partial x_1} & -\frac{\partial A_3}{\partial x_1} \\ \frac{\partial \phi}{\partial x_2} & -\frac{\partial A_1}{\partial x_2} & -\frac{\partial A_2}{\partial x_2} & -\frac{\partial A_3}{\partial x_2} \\ \frac{\partial \phi}{\partial x_3} & -\frac{\partial A_1}{\partial x_3} & -\frac{\partial A_2}{\partial x_3} & -\frac{\partial A_3}{\partial x_3} \end{pmatrix} , \tag{2.30}$$

thus

$$\partial^{\mathrm{T}} A^{\mathrm{T}} \eta - \left(\partial^{\mathrm{T}} A^{\mathrm{T}} \eta\right)^{\mathrm{T}} = \begin{pmatrix} 0 & -\frac{\partial \phi}{\partial x_1} - \frac{1}{c}\frac{\partial A_1}{\partial t} & -\frac{\partial \phi}{\partial x_2} - \frac{1}{c}\frac{\partial A_2}{\partial t} & -\frac{\partial \phi}{\partial x_3} - \frac{1}{c}\frac{\partial A_3}{\partial t} \\ \frac{\partial \phi}{\partial x_1} + \frac{1}{c}\frac{\partial A_1}{\partial t} & 0 & \frac{\partial A_1}{\partial x_2} - \frac{\partial A_2}{\partial x_1} & \frac{\partial A_1}{\partial x_3} - \frac{\partial A_3}{\partial x_1} \\ \frac{\partial \phi}{\partial x_2} + \frac{1}{c}\frac{\partial A_2}{\partial t} & \frac{\partial A_2}{\partial x_1} - \frac{\partial A_1}{\partial x_2} & 0 & \frac{\partial A_2}{\partial x_3} - \frac{\partial A_3}{\partial x_2} \\ \frac{\partial \phi}{\partial x_3} + \frac{1}{c}\frac{\partial A_3}{\partial t} & \frac{\partial A_3}{\partial x_1} - \frac{\partial A_1}{\partial x_3} & \frac{\partial A_3}{\partial x_2} - \frac{\partial A_2}{\partial x_3} & 0 \end{pmatrix} , \tag{2.31}$$

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 A, implies the following:

Claim. The field matrix $\hat{F}$ is Lorentz transformed according to

$$\hat{F} = \Lambda^{\mathrm{T}} \hat{F}' \Lambda \,, \tag{2.32}$$

provided that A is Lorentz transformed according to $A' = \Lambda A$.

Proof. The assumption $A' = \Lambda A$ leads to [recall Eq. (1.115), which reads $\partial' = \partial \Lambda^{-1}$ and Eq. (1.21), which reads $\Lambda^{-1} = \eta \Lambda^{\mathrm{T}} \eta$]

$$\begin{aligned} \partial^{\mathrm{T}} A^{\mathrm{T}} \eta &= \Lambda^{\mathrm{T}} \left(\partial'\right)^{\mathrm{T}} \left(A'\right)^{\mathrm{T}} \left(\Lambda^{-1}\right)^{\mathrm{T}} \eta \\ &= \Lambda^{\mathrm{T}} \left(\left(\partial'\right)^{\mathrm{T}} \left(A'\right)^{\mathrm{T}} \eta \right) \Lambda \,, \end{aligned} \tag{2.33}$$

and thus

$$\partial^{\mathrm{T}} A^{\mathrm{T}} \eta - \left(\partial^{\mathrm{T}} A^{\mathrm{T}} \eta\right)^{\mathrm{T}} = \Lambda^{\mathrm{T}} \left[\left(\left(\partial'\right)^{\mathrm{T}} \left(A'\right)^{\mathrm{T}} \eta \right) - \left(\left(\left(\partial'\right)^{\mathrm{T}} \left(A'\right)^{\mathrm{T}} \eta \right) \right)^{\mathrm{T}} \right] \Lambda \,, \tag{2.34}$$

in agreement with Eq. (2.32).

Exercise 2.2.1. Let $\mathbf{u}$ be the relative velocity of frame S' with respect to frame S. Show using Eq. (2.32) that the fields $\mathbf{E}$ and $\mathbf{B}$ are transformed according to

$$\mathbf{E} = \mathbf{E}'_{\parallel} + \gamma \left(\mathbf{E}'_{\perp} - \boldsymbol{\beta} \times \mathbf{B}'_{\perp} \right) \,, \tag{2.35}$$

$$\mathbf{B} = \mathbf{B}'_{\parallel} + \gamma \left(\mathbf{B}'_{\perp} + \boldsymbol{\beta} \times \mathbf{E}'_{\perp} \right) \,, \tag{2.36}$$

where $\mathbf{V}_{\parallel}$ ($\mathbf{V}_{\perp}$) denotes the component of a 3-vector $\mathbf{V}$ parallel (perpendicular) to $\mathbf{u}$ and where $\boldsymbol{\beta} = \mathbf{u}/c$.

Solution 2.2.1. When the coordinates of frame S are chosen such that the velocity $\mathbf{u}$ is pointing in the x_1 direction Λ becomes [see Eq. (1.26)]

$$\Lambda = \begin{pmatrix} \gamma & -\gamma\beta & 0 & 0 \\ -\gamma\beta & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \,, \tag{2.37}$$

where $\beta = u/c$ and $\gamma = 1/\sqrt{1-\beta^2}$. The transformation (2.32) yields

$$E_1 = E'_1 \,, \tag{2.38}$$

$$E_2 = \gamma \left(E'_2 + \beta B'_3 \right) \,, \tag{2.39}$$

$$E_3 = \gamma \left(E'_3 - \beta B'_2 \right) \,, \tag{2.40}$$

$$B_1 = B'_1 \,, \tag{2.41}$$

$$B_2 = \gamma \left(B'_2 - \beta E'_3 \right) \,, \tag{2.42}$$

$$B_3 = \gamma \left(B'_3 + \beta E'_2 \right) \,. \tag{2.43}$$

In a vectorial form the above can be rewritten as Eqs. (2.35) and (2.36).

Note that for the case where $\mathbf{B}'$ vanishes Eqs. (2.35) and (2.36) coincide with Eqs. (1.108) and (1.109).

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

$$A' = \Lambda A \,. \tag{2.44}$$

Claim. The inhomogeneous Maxwell's equations (2.12) and (2.14) can be written as [see Eq. (1.114)]

$$\partial \eta \hat{G} \eta = (4\pi/c)\, J_{\text{ext}}^{\text{T}} \,, \tag{2.45}$$

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

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

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

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

and thus [see Eq. (1.114)]

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

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)]

$$\hat{G} = \Lambda^{\text{T}} \hat{G}' \Lambda \,. \tag{2.49}$$

Proof. With the help of Eq. (1.21), which reads $\Lambda^{-1} = \eta \Lambda^{\text{T}} \eta$, Eq. (1.113), which for the case of J_{ext} becomes $J'_{\text{ext}} = \Lambda J_{\text{ext}}$, Eq. (1.115), which reads $\partial' = \partial \Lambda^{-1}$, and Eq. (2.49), which reads $\hat{G} = \Lambda^{\text{T}} \hat{G}' \Lambda$, one finds that

$$\begin{aligned} \partial \eta \hat{G} \eta &= \partial \eta \Lambda^{\text{T}} \hat{G}' \Lambda \eta \\ &= \partial' \Lambda \eta \Lambda^{\text{T}} \eta \eta \hat{G}' \Lambda \eta \\ &= \partial' \eta \hat{G}' \Lambda \eta \\ &= \partial' \eta \hat{G}' \eta \left(\eta \Lambda^{\text{T}} \eta \right)^{\text{T}} \\ &= \partial' \eta \hat{G}' \eta \left(\Lambda^{-1} \right)^{\text{T}} \,, \end{aligned} \tag{2.50}$$

and

$$J_{\mathrm{ext}}^{\mathrm{T}} = \left(\Lambda^{-1} J'_{\mathrm{ext}}\right)^{\mathrm{T}} = J'^{\mathrm{T}}_{\mathrm{ext}} \left(\Lambda^{-1}\right)^{\mathrm{T}} , \tag{2.51}$$

and thus

$$\partial' \eta \hat{G}' \eta = \frac{4\pi}{c} J'^{\mathrm{T}}_{\mathrm{ext}} , \tag{2.52}$$

i.e. the relation (2.45) is Lorentz invariant.

2.4 Gauge Transformation

The relation between the fields **E** and **B** and the scalar ϕ and the 3-vector **A** potentials is given by Eqs. (2.23) and (2.25). For given fields **E** and **B**, however, the potentials ϕ and **A** are not uniquely determined by Eqs. (2.23) and (2.25), as can be demonstrated by the following transformation

$$A \rightarrow A' = A + (\partial \psi \eta)^{\mathrm{T}} , \tag{2.53}$$

or [see Eqs. (1.114) and (2.26)]

$$\phi \rightarrow \phi' = \phi + \frac{1}{c}\frac{\partial \psi}{\partial t} , \tag{2.54}$$

$$\mathbf{A} \rightarrow \mathbf{A}' = \mathbf{A} - \boldsymbol{\nabla} \psi , \tag{2.55}$$

where $\psi(t, \mathbf{r})$ 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 (1.118) and (1.120) in terms of the potentials ϕ and **A**.

Solution 2.5.1. Substituting Eqs. (2.23) and (2.25) into the Maxwell's equations in vacuum (1.118) and (1.120) leads with the help of the vector identity

$$\boldsymbol{\nabla} \times (\boldsymbol{\nabla} \times \mathbf{A}) = \boldsymbol{\nabla} (\boldsymbol{\nabla} \cdot \mathbf{A}) - \boldsymbol{\nabla}^2 \mathbf{A} , \tag{2.56}$$

to

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

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

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

$$\Box^2 = -\frac{1}{c^2}\frac{\partial^2}{\partial t^2} + \boldsymbol{\nabla}^2 \;, \tag{2.59}$$

is the D'Alembertian operator, and where ∂A is given by [see Eq. (2.27)].

$$\partial A = \frac{1}{c}\frac{\partial \phi}{\partial t} + \boldsymbol{\nabla} \cdot \mathbf{A} \;. \tag{2.60}$$

2.5.1 Lorenz Gauge

The choice, for which

$$0 = \frac{1}{c}\frac{\partial \phi}{\partial t} + \boldsymbol{\nabla} \cdot \mathbf{A} \;, \tag{2.61}$$

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 ∂A. For this case the Maxwell's equations in vacuum (2.57) and (2.58) become

$$\Box^2 \mathbf{A} = -\frac{4\pi}{c}\mathbf{J} \;, \tag{2.62}$$

$$\Box^2 \phi = -4\pi\rho \;. \tag{2.63}$$

In electrostatics the Poisson's equation (1.5), which is given by $\boldsymbol{\nabla}^2\phi = -4\pi\rho$, relates the 0'th component of the potential 4-vector $A = (\phi, A_1, A_2, A_3)^{\mathrm{T}}$ with the 0'th component of the current 4-vector $J = (c\rho, J_1, J_2, J_3)^{\mathrm{T}}$. 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

$$0 = \boldsymbol{\nabla} \cdot \mathbf{A} \;. \tag{2.64}$$

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

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

$$\boldsymbol{\nabla}^2\phi = -4\pi\rho \;. \tag{2.66}$$

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

$$\int_S (\boldsymbol{\nabla} \times \mathbf{V}) \cdot \mathrm{d}\mathbf{s} = \oint_C \mathbf{V} \cdot \mathrm{d}\mathbf{l} \,, \tag{2.67}$$

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

$$\int_V (\boldsymbol{\nabla} \cdot \mathbf{V}) \; \mathrm{d}v = \int_S \mathbf{V} \cdot \mathrm{d}\mathbf{s} \,, \tag{2.68}$$

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

$$\oint_C \mathbf{H} \cdot \mathrm{d}\mathbf{l} = \frac{4\pi}{c} \int_S (\mathbf{J}_{\mathrm{ext}} + \mathbf{J}_{\mathrm{cond}}) \cdot \mathrm{d}\mathbf{s} + \frac{1}{c}\frac{\partial}{\partial t} \int_S \mathbf{D} \cdot \mathrm{d}\mathbf{s} \,, \tag{2.69}$$

$$\oint_C \mathbf{E} \cdot \mathrm{d}\mathbf{l} = -\frac{1}{c}\frac{\partial}{\partial t} \int_S \mathbf{B} \cdot \mathrm{d}\mathbf{s} \,, \tag{2.70}$$

$$\int_S \mathbf{D} \cdot \mathrm{d}\mathbf{s} = 4\pi \int_V \rho_{\mathrm{ext}} \; \mathrm{d}v \,, \tag{2.71}$$

$$\int_S \mathbf{B} \cdot \mathrm{d}\mathbf{s} = 0 \,. \tag{2.72}$$

Consider an interface between two materials. Let ρ_{s} be the areal charge density and let $\mathbf{J}_{\mathrm{s}}$ be the surface current density on the boundary surface. Let $\hat{\mathbf{n}}$ 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 (1.96), which is given by

$$\mathbf{A} \times (\mathbf{B} \times \mathbf{C}) = (\mathbf{A} \cdot \mathbf{C})\,\mathbf{B} - (\mathbf{A} \cdot \mathbf{B})\,\mathbf{C} \,, \tag{2.73}$$

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

$$\mathbf{V} = \mathbf{V}_{\mathrm{n}} + \mathbf{V}_{\mathrm{t}} \,, \tag{2.74}$$

where

$$\mathbf{V}_{\mathrm{n}} = \hat{\mathbf{n}}\,(\hat{\mathbf{n}} \cdot \mathbf{V}) \;, \tag{2.75}$$

$$\mathbf{V}_{\mathrm{t}} = \hat{\mathbf{n}} \times (\mathbf{V} \times \hat{\mathbf{n}}) \;. \tag{2.76}$$

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)

$$\hat{\mathbf{n}} \times (\mathbf{H}_2 - \mathbf{H}_1) = \frac{4\pi}{c}\mathbf{J}_\mathrm{s} , \tag{2.77}$$

$$\hat{\mathbf{n}} \times (\mathbf{E}_2 - \mathbf{E}_1) = 0 , \tag{2.78}$$

$$\hat{\mathbf{n}} \cdot (\mathbf{D}_2 - \mathbf{D}_1) = 4\pi\rho_\mathrm{s} , \tag{2.79}$$

$$\hat{\mathbf{n}} \cdot (\mathbf{B}_2 - \mathbf{B}_1) = 0 . \tag{2.80}$$

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

Solution 2.6.1. Inside a perfect conductor ($\sigma \to \infty$) all fields vanish, and consequently the boundary conditions (2.77), (2.78), (2.79) and (2.80) become

$$\hat{\mathbf{n}} \times \mathbf{H} = \frac{4\pi}{c}\mathbf{J}_\mathrm{s} , \tag{2.81}$$

$$\hat{\mathbf{n}} \times \mathbf{E} = 0 , \tag{2.82}$$

$$\hat{\mathbf{n}} \cdot \mathbf{D} = 4\pi\rho_\mathrm{s} , \tag{2.83}$$

$$\hat{\mathbf{n}} \cdot \mathbf{B} = 0 . \tag{2.84}$$

2.7 Isotropic and Linear Medium

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

$$\mathbf{D} = \epsilon\mathbf{E} , \tag{2.85}$$

$$\mathbf{P} = \chi_\mathrm{e}\mathbf{E} , \tag{2.86}$$

$$\mathbf{B} = \mu\mathbf{H} , \tag{2.87}$$

$$\mathbf{M} = \chi_\mathrm{m}\mathbf{H} , \tag{2.88}$$

where ϵ is the relative permittivity, χ_e is the electric susceptibility, μ is the relative permeability and χ_m is the magnetic susceptibility, where [see Eqs. (2.17) and (2.18)]

$$\epsilon = 1 + 4\pi\chi_\mathrm{e} , \tag{2.89}$$

$$\mu = 1 + 4\pi\chi_\mathrm{m} . \tag{2.90}$$

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

$$\mathbf{J}_\mathrm{cond} = \sigma\mathbf{E} , \tag{2.91}$$

where σ is the conductivity.

Exercise 2.7.1. energy conservation - Show that for the case where $\mathbf{J}_\mathrm{ext} = 0$ the following holds

$$\int_S \mathbf{S} \cdot \mathrm{d}\mathbf{s} + \int_V \sigma\mathbf{E}^2 \,\mathrm{d}v + \frac{\partial}{\partial t}\int_V u \,\mathrm{d}v = 0 , \tag{2.92}$$

where $\mathbf{S}$ is the Poynting vector, which is given by

$$\mathbf{S} = \frac{c}{4\pi}\mathbf{E}\times\mathbf{H} \, , \tag{2.93}$$

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

$$u = \frac{\epsilon\mathbf{E}^2 + \mu\mathbf{H}^2}{8\pi} \, . \tag{2.94}$$

Solution 2.7.1. By multiplying Eq. (2.12) by $\mathbf{E}$, multiplying Eq. (2.13) by $\mathbf{H}$, subtracting and employing the vector identity (1.132), which is given by

$$\boldsymbol{\nabla}\cdot(\mathbf{A}\times\mathbf{B}) = \mathbf{B}\cdot(\boldsymbol{\nabla}\times\mathbf{A}) - \mathbf{A}\cdot(\boldsymbol{\nabla}\times\mathbf{B}) \, , \tag{2.95}$$

one obtains

$$\nabla\cdot\mathbf{S} + \mathbf{E}\cdot\mathbf{J}_{\mathrm{cond}} + \frac{1}{4\pi}\left(\mathbf{E}\cdot\frac{\partial\mathbf{D}}{\partial t} + \mathbf{H}\cdot\frac{\partial\mathbf{B}}{\partial t}\right) = 0 \, , \tag{2.96}$$

or in terms of u [see Eq. (2.91)]

$$\nabla\cdot\mathbf{S} + \sigma\mathbf{E}^2 + \frac{\partial u}{\partial t} = 0 \, . \tag{2.97}$$

Applying the divergence theorem (2.68) leads to Eq. (2.92).

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

$$\mathbf{D} + \boldsymbol{\beta}\times\mathbf{H} = \epsilon\left(\mathbf{E} + \boldsymbol{\beta}\times\mathbf{B}\right) \, , \tag{2.98}$$

$$\mathbf{B} - \boldsymbol{\beta}\times\mathbf{E} = \mu\left(\mathbf{H} - \boldsymbol{\beta}\times\mathbf{D}\right) \, , \tag{2.99}$$

where $\boldsymbol{\beta} = \mathbf{u}/c$.

Solution 2.7.2. Let S' be the rest frame of the medium, in which the following constitutive relations hold [see Eqs. (2.85) and (2.87)]

$$\mathbf{D}' = \epsilon\mathbf{E}' \, , \tag{2.100}$$

$$\mathbf{B}' = \mu\mathbf{H}' \, . \tag{2.101}$$

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

$$\mathbf{E}' = \mathbf{E}_{\|} + \gamma\left(\mathbf{E}_{\perp} + \boldsymbol{\beta}\times\mathbf{B}_{\perp}\right) \, , \tag{2.102}$$

$$\mathbf{B}' = \mathbf{B}_{\|} + \gamma\left(\mathbf{B}_{\perp} - \boldsymbol{\beta}\times\mathbf{E}_{\perp}\right) \, , \tag{2.103}$$

and [see Eq. (2.49)]

$$\mathbf{D}' = \mathbf{D}_{\|} + \gamma\left(\mathbf{D}_{\perp} + \boldsymbol{\beta}\times\mathbf{H}_{\perp}\right) \, , \tag{2.104}$$

$$\mathbf{H}' = \mathbf{H}_{\|} + \gamma\left(\mathbf{H}_{\perp} - \boldsymbol{\beta}\times\mathbf{D}_{\perp}\right) \, , \tag{2.105}$$

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

$$\mathbf{D}_{\parallel} + \gamma\left(\mathbf{D}_{\perp} + \boldsymbol{\beta} \times \mathbf{H}_{\perp}\right) = \epsilon\left[\mathbf{E}_{\parallel} + \gamma\left(\mathbf{E}_{\perp} + \boldsymbol{\beta} \times \mathbf{B}_{\perp}\right)\right] , \tag{2.106}$$

$$\mathbf{B}_{\parallel} + \gamma\left(\mathbf{B}_{\perp} - \boldsymbol{\beta} \times \mathbf{E}_{\perp}\right) = \mu\left[\mathbf{H}_{\parallel} + \gamma\left(\mathbf{H}_{\perp} - \boldsymbol{\beta} \times \mathbf{D}_{\perp}\right)\right] . \tag{2.107}$$

By dividing the perpendicular components of both sides of both Eqs. (2.106) and (2.107) by γ one obtains

$$\mathbf{D}_{\parallel} + \mathbf{D}_{\perp} + \boldsymbol{\beta} \times \mathbf{H}_{\perp} = \epsilon\left(\mathbf{E}_{\parallel} + \mathbf{E}_{\perp} + \boldsymbol{\beta} \times \mathbf{B}_{\perp}\right) , \tag{2.108}$$

$$\mathbf{B}_{\parallel} + \mathbf{B}_{\perp} - \boldsymbol{\beta} \times \mathbf{E}_{\perp} = \mu\left(\mathbf{H}_{\parallel} + \mathbf{H}_{\perp} - \boldsymbol{\beta} \times \mathbf{D}_{\perp}\right) , \tag{2.109}$$

and thus Eqs. (2.98) and (2.99) hold (note that $\mathbf{V}_{\parallel} + \mathbf{V}_{\perp} = \mathbf{V}$ and $\boldsymbol{\beta} \times \mathbf{V}_{\perp} = \boldsymbol{\beta} \times \mathbf{V}$).

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

$$\partial g \hat{F} g = \frac{4\pi}{c} J_{\mathrm{ext}}^{\mathrm{T}} , \tag{2.110}$$

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

$$\hat{F} = \begin{pmatrix} 0 & E_1 & E_2 & E_3 \\ -E_1 & 0 & -B_3 & B_2 \\ -E_2 & B_3 & 0 & -B_1 \\ -E_3 & -B_2 & B_1 & 0 \end{pmatrix} , \tag{2.111}$$

the effective metric g is given by

$$g = \frac{1}{\sqrt{\mu}}\left(\eta + \frac{\xi}{c^2} U U^{\mathrm{T}}\right) , \tag{2.112}$$

η is the Minkowski metric (1.14), the velocity 4-vector U is defined by [compare with Eq. (1.64)]

$$U = \frac{\mathrm{d}X}{\mathrm{d}\tau} = \gamma\left(c, u_1, u_2, u_3\right)^{\mathrm{T}} . \tag{2.113}$$

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

$$\xi = \epsilon\mu - 1 , \tag{2.114}$$

ϵ is the relative permittivity, μ is the relative permeability and the 4-vector $J_{\mathrm{ext}} = \left(c\rho_{\mathrm{ext}}, \mathbf{J}_{\mathrm{ext}} + \mathbf{J}_{\mathrm{cond}}\right)^{\mathrm{T}}$ is defined by Eq. (2.22).

Proof. Let S' be the rest frame of the medium. The constitutive relations $\mathbf{D}' = \epsilon \mathbf{E}'$ (2.85) and $\mathbf{B}' = \mu \mathbf{H}'$ (2.87) can be expressed as [see Eqs. (2.29) and (2.46)]

$$\hat{G}' = \zeta \hat{F}' \zeta \ , \tag{2.115}$$

where

$$\hat{F}' = \begin{pmatrix} 0 & E_1' & E_2' & E_3' \\ -E_1' & 0 & -B_3' & B_2' \\ -E_2' & B_3' & 0 & -B_1' \\ -E_3' & -B_2' & B_1' & 0 \end{pmatrix} \ , \tag{2.116}$$

$$\hat{G}' = \begin{pmatrix} 0 & D_1' & D_2' & D_3' \\ -D_1' & 0 & -H_3' & H_2' \\ -D_2' & H_3' & 0 & -H_1' \\ -D_3' & -H_2' & H_1' & 0 \end{pmatrix} \ , \tag{2.117}$$

and where the matrix ζ is given by

$$\zeta = \frac{1}{\sqrt{\mu}} \begin{pmatrix} 1+\xi & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \ , \tag{2.118}$$

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

$$\hat{F}' = \left(\Lambda^{-1}\right)^{\mathrm{T}} \hat{F} \Lambda^{-1} \ , \tag{2.119}$$

$$\hat{G}' = \left(\Lambda^{-1}\right)^{\mathrm{T}} \hat{G} \Lambda^{-1} \ , \tag{2.120}$$

and thus Eq. (2.115) can be rewritten as

$$\hat{G} = \Lambda^{\mathrm{T}} \zeta \left(\Lambda^{-1}\right)^{\mathrm{T}} \hat{F} \Lambda^{-1} \zeta \Lambda \ . \tag{2.121}$$

The above result (2.121) allows writing the Maxwell's equation $\partial \eta \hat{G} \eta = (4\pi/c) J_{\mathrm{ext}}^{\mathrm{T}}$ (2.45) as

$$\partial g^{\mathrm{T}} \hat{F} g = \frac{4\pi}{c} J_{\mathrm{ext}}^{\mathrm{T}} \ , \tag{2.122}$$

where

$$g = \Lambda^{-1} \zeta \Lambda \eta \ . \tag{2.123}$$

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

$$\Lambda^{-1} \zeta \Lambda \eta = \frac{1}{\sqrt{\mu}} \left(\eta + \frac{\xi}{c^2} U U^{\mathrm{T}} \right) \ , \tag{2.124}$$

in agreement with Eq. (2.112). Note that $g^{\mathrm{T}} = g$. In a matrix form the metric g (2.112) is given by

$$g = \frac{1}{\sqrt{\mu}} \left(\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} + \frac{\xi}{1-\beta^2} \begin{pmatrix} 1 & \beta_1 & \beta_2 & \beta_3 \\ \beta_1 & \beta_1^2 & \beta_1\beta_2 & \beta_1\beta_3 \\ \beta_2 & \beta_2\beta_1 & \beta_2^2 & \beta_2\beta_3 \\ \beta_3 & \beta_3\beta_1 & \beta_3\beta_2 & \beta_3^2 \end{pmatrix} \right) , \tag{2.125}$$

where $\boldsymbol{\beta}$ is related to the velocity 3-vector $\mathbf{u}$ by $\boldsymbol{\beta} = \mathbf{u}/c$ [see Eq. (2.113)].

With the help of the relation $\partial^{\mathrm{T}} A^{\mathrm{T}} \eta - \left(\partial^{\mathrm{T}} A^{\mathrm{T}} \eta\right)^{\mathrm{T}} = \hat{F}$ [see Eq. (2.28)] Eq. (2.110) can be expressed in terms of the potential 4-vector A as

$$\partial g \left(\partial^{\mathrm{T}} A^{\mathrm{T}} \eta - \left(\partial^{\mathrm{T}} A^{\mathrm{T}} \eta\right)^{\mathrm{T}} \right) g = \frac{4\pi}{c} J_{\mathrm{ext}}^{\mathrm{T}} . \tag{2.126}$$

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 ω. It is convenient to employ complex notation, in which all fields and sources are expressed as

$$\psi(\mathbf{r}, t) = \mathrm{real}\left[\psi(\mathbf{r})\, e^{-i\omega t}\right] . \tag{2.127}$$

Note that in order to avoid cumbersome notation, the same letter ψ denotes the $\mathbf{r}$ and t dependent amplitude $\psi(\mathbf{r}, t)$ and the $\mathbf{r}$ only dependent amplitude $\psi(\mathbf{r})$, 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

$$\boldsymbol{\nabla} \times \mathbf{H} = \frac{4\pi \mathbf{J}_{\mathrm{ext}}}{c} - \frac{i\omega}{c} \epsilon_{\mathrm{eff}} \mathbf{E} , \tag{2.128}$$

$$\boldsymbol{\nabla} \times \mathbf{E} = \frac{i\omega}{c} \mu \mathbf{H} , \tag{2.129}$$

$$\boldsymbol{\nabla} \cdot (\epsilon \mathbf{E}) = 4\pi \rho_{\mathrm{ext}} , \tag{2.130}$$

$$\boldsymbol{\nabla} \cdot (\mu \mathbf{H}) = 0 , \tag{2.131}$$

where the effective relative dielectric coefficient ϵ_{eff} is given by

$$\epsilon_{\mathrm{eff}} = \epsilon + i \frac{4\pi\sigma}{\omega} . \tag{2.132}$$

Exercise 2.8.1. Consider two vectors $\mathbf{V}_{\mathrm{a}}$ and $\mathbf{V}_{\mathrm{b}}$ having harmonic time dependency

$$\mathbf{V}_{\mathrm{a}}(\mathbf{r}, t) = \mathrm{real}\left[\mathbf{V}_{\mathrm{a}}(\mathbf{r})\, e^{-i\omega t}\right] , \tag{2.133}$$

$$\mathbf{V}_{\mathrm{b}}(\mathbf{r}, t) = \mathrm{real}\left[\mathbf{V}_{\mathrm{b}}(\mathbf{r})\, e^{-i\omega t}\right] . \tag{2.134}$$

Calculate the time averaged of $\mathbf{V}_{\mathrm{a}}(\mathbf{r}, t) \cdot \mathbf{V}_{\mathrm{b}}(\mathbf{r}, t)$.

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

$$
\begin{aligned}
\left\langle \mathbf{V}_{\mathrm{a}}(\mathbf{r},t)\cdot\mathbf{V}_{\mathrm{b}}(\mathbf{r},t)\right\rangle &= \sum_{n=1}^{3}\left\langle \mathrm{real}\left[V_{\mathrm{a},n}e^{-i\omega t}\right]\mathrm{real}\left[V_{\mathrm{b},n}e^{-i\omega t}\right]\right\rangle \\
&= \sum_{n=1}^{3}\left\langle \frac{V_{\mathrm{a},n}e^{-i\omega t}+V_{\mathrm{a},n}^{*}e^{i\omega t}}{2}\frac{V_{\mathrm{b},n}e^{-i\omega t}+V_{\mathrm{b},n}^{*}e^{i\omega t}}{2}\right\rangle \\
&= \frac{1}{2}\sum_{n=1}^{3}\mathrm{real}\left(V_{\mathrm{a},n}V_{\mathrm{b},n}^{*}\right)\ , \qquad (2.135) \\
& \qquad (2.136)
\end{aligned}
$$

thus

$$
\left\langle \mathbf{V}_{\mathrm{a}}(\mathbf{r},t)\cdot\mathbf{V}_{\mathrm{b}}(\mathbf{r},t)\right\rangle = \frac{1}{2}\mathrm{real}\left(\mathbf{V}_{\mathrm{a}}\cdot\mathbf{V}_{\mathrm{b}}^{*}\right) \qquad (2.137)
$$

2.9 Inhomogeneous Medium Free of Sources

Consider the case of electromagnetic fields in an inhomogeneous medium free of sources (i.e. $\rho_{\mathrm{ext}} = 0$ and $\mathbf{J}_{\mathrm{ext}} = 0$,), which is assumed to be isotropic, linear and stationary. In addition, the conductivity σ 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. (2.128), (2.129), (2.130) and (2.131) become

$$
\begin{aligned}
\boldsymbol{\nabla}\times\mathbf{H} &= -ik_0\epsilon\mathbf{E}\ , \qquad (2.138) \\
\boldsymbol{\nabla}\times\mathbf{E} &= ik_0\mu\mathbf{H}\ , \qquad (2.139) \\
\boldsymbol{\nabla}\cdot(\epsilon\mathbf{E}) &= 0\ , \qquad (2.140) \\
\boldsymbol{\nabla}\cdot(\mu\mathbf{H}) &= 0\ , \qquad (2.141)
\end{aligned}
$$

where

$$
k_0 = \frac{\omega}{c}\ . \qquad (2.142)
$$

Note that Eqs. (2.140) and (2.141) result from Eqs. (2.138) and (2.139) by the vector identity

$$
\nabla\cdot(\nabla\times\mathbf{A}) = 0\ . \qquad (2.143)
$$

Exercise 2.9.1. Show that

$$
\begin{aligned}
\boldsymbol{\nabla}^2\mathbf{E}+n^2k_0^2\mathbf{E}+(\boldsymbol{\nabla}\log\mu)\times(\boldsymbol{\nabla}\times\mathbf{E})+\boldsymbol{\nabla}(\mathbf{E}\cdot\boldsymbol{\nabla}\log\epsilon) &= 0\ , \qquad (2.144) \\
\boldsymbol{\nabla}^2\mathbf{H}+n^2k_0^2\mathbf{H}+(\boldsymbol{\nabla}\log\epsilon)\times(\boldsymbol{\nabla}\times\mathbf{H})+\boldsymbol{\nabla}(\mathbf{H}\cdot\boldsymbol{\nabla}\log\mu) &= 0\ , \qquad (2.145)
\end{aligned}
$$

where n, which is given by

$$n = \sqrt{\epsilon\mu} \ , \tag{2.146}$$

is the so-called refraction index.

Solution 2.9.1. Applying the operator $\boldsymbol{\nabla}\times$ to Eq. (2.138) and using the vector identity (2.56), which is given by

$$\boldsymbol{\nabla}\times(\boldsymbol{\nabla}\times\mathbf{A}) = \boldsymbol{\nabla}(\boldsymbol{\nabla}\cdot\mathbf{A}) - \boldsymbol{\nabla}^2\mathbf{A} \ , \tag{2.147}$$

one finds that

$$\boldsymbol{\nabla}(\boldsymbol{\nabla}\cdot\mathbf{H}) - \boldsymbol{\nabla}^2\mathbf{H} = -\frac{i\omega}{c}\boldsymbol{\nabla}\times(\epsilon\mathbf{E}) \ . \tag{2.148}$$

Using the vectors identities

$$\boldsymbol{\nabla}\cdot(f\mathbf{A}) = f\boldsymbol{\nabla}\cdot\mathbf{A} + \mathbf{A}\cdot\boldsymbol{\nabla}f \ , \tag{2.149}$$

and

$$\boldsymbol{\nabla}\times(f\mathbf{A}) = f\boldsymbol{\nabla}\times\mathbf{A} + (\boldsymbol{\nabla}f)\times\mathbf{A} \ , \tag{2.150}$$

together with Eqs. (2.139) and (2.141) leads to Eq. (2.145). Equation (2.144) 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. (2.144) and (2.145) are disregarded. These terms give rise to coupling between the components of **E** and **H**. After the removal of these terms Eqs. (2.144) and (2.145) imply that all three components of **E** and **H** satisfy the so-called Helmholtz equation, which is given by

$$\left(\boldsymbol{\nabla}^2 + n^2k_0^2\right)\psi = 0 \ . \tag{2.151}$$

Note that for the case of a homogeneous medium, in which both ϵ and μ are constants, the above statement [i.e. all three components of **E** and **H** satisfy the Helmholtz equation (2.151)] becomes exact.

2.11 Polarization

Consider an electric field, which is denoted in this section by $\mathbf{E}_{\mathrm{t}}$, having harmonic time dependency

$$\mathbf{E}_{\mathrm{t}} = \mathrm{real}\left[\exp(-i\omega t)\,\mathbf{E}\right] \ , \tag{2.152}$$

where the complex phasor vector $\mathbf{E}$ is taken to be given by

$$\mathbf{E} = \mathbf{a} + i\mathbf{b} \,, \tag{2.153}$$

where $\mathbf{a}$ and $\mathbf{b}$ are real vectors (constants for a given spacial position). The real time dependent field is given by

$$\begin{aligned}\mathbf{E}_\mathrm{t} &= \frac{1}{2}\left[\exp\left(-i\omega t\right)\mathbf{E} + \exp\left(i\omega t\right)\mathbf{E}^*\right] \\ &= \frac{1}{2}\left[\exp\left(-i\omega t\right)\left(\mathbf{a}+i\mathbf{b}\right) + \exp\left(i\omega t\right)\left(\mathbf{a}-i\mathbf{b}\right)\right] \\ &= \mathbf{a}\cos\left(\omega t\right) + \mathbf{b}\sin\left(\omega t\right) \ .\end{aligned} \tag{2.154}$$

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

Claim. The curve $\mathbf{E}_\mathrm{t}\left(t\right)$ is an ellipse.

Proof. To show that $\mathbf{E}_\mathrm{t}\left(t\right) = \left(E_x, E_y, E_z\right)$ 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

$$\sum_{0\leq n_x+n_y+n_z\leq 2} A_{n_x,n_y,n_z} E_x^{n_x} E_y^{n_y} E_z^{n_z} = 0 \,, \tag{2.155}$$

where all coefficients A_{n_x,n_y,n_z} are time independent real constants. The following holds [see Eq. (2.154)]

$$E_i^2\left(t\right) = a_i^2\cos^2\left(\omega t\right) + b_i^2\sin^2\left(\omega t\right) + a_i b_i \sin\left(2\omega t\right) \ , \tag{2.156}$$

or

$$E_i^2\left(t\right) - \frac{1}{2}\left(a_i^2+b_i^2\right) = \frac{1}{2}\left(a_i^2-b_i^2\right)\cos\left(2\omega t\right) + a_i b_i \sin\left(2\omega t\right) \ , \tag{2.157}$$

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

$$M\begin{pmatrix}\cos\left(2\omega t\right)\\ \sin\left(2\omega t\right)\\ -1\end{pmatrix} = 0 \,, \tag{2.158}$$

where

$$M = \begin{pmatrix}\frac{a_x^2-b_x^2}{2} & a_x b_x & E_x^2\left(t\right) - \frac{a_x^2+b_x^2}{2}\\ \frac{a_y^2-b_y^2}{2} & a_y b_y & E_y^2\left(t\right) - \frac{a_y^2+b_y^2}{2}\\ \frac{a_z^2-b_z^2}{2} & a_z b_z & E_z^2\left(t\right) - \frac{a_z^2+b_z^2}{2}\end{pmatrix} \ . \tag{2.159}$$

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

$$\det M = 0 \,, \tag{2.160}$$

thus, $\mathbf{E}_\mathrm{t}\left(t\right)$ is indeed conic section and thus (since it is periodic) an ellipse.

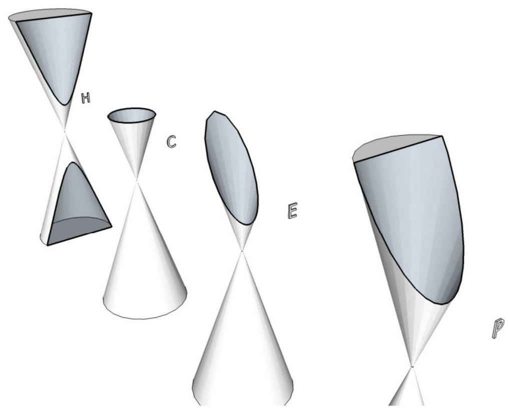

Fig. 2.1. Conic sections: hyperbole (H), circle (C), ellipse (E) and parabola (P).

The eccentricity e of an ellipse is defined by

$$e=\sqrt{1-\frac{(\mathbf{E}_{\mathrm{t}}^2)_{\min}}{(\mathbf{E}_{\mathrm{t}}^2)_{\max}}}, \tag{2.161}$$

where $\left(\mathbf{E}_{\mathrm{t}}^2\right)_{\min}$ ($\left(\mathbf{E}_{\mathrm{t}}^2\right)_{\max}$) is the minimum (maximum) value of $\mathbf{E}_{\mathrm{t}}^2$.

Exercise 2.11.1. Show that eccentricity is given by

$$e=\sqrt{\frac{2\left|\eta\right|}{1+\left|\eta\right|}}\ , \tag{2.162}$$

where

$$\eta=\frac{\mathbf{E}^2}{\left|\mathbf{E}\right|^2}\ . \tag{2.163}$$

Solution 2.11.1. Taking the square of $\mathbf{E}_{\mathrm{t}}\left(t\right)$ leads to [see Eq. (2.152)]

$$\mathbf{E}_{\mathrm{t}}^2\left(t\right)=\frac{1}{4}\left[\exp\left(-2i\omega t\right)\mathbf{E}^2+\exp\left(2i\omega t\right)\left(\mathbf{E}^*\right)^2+2\mathbf{E}\cdot\mathbf{E}^*\right]\ . \tag{2.164}$$

Using the notation

$$\eta=\frac{\mathbf{E}^2}{\left|\mathbf{E}\right|^2}\ , \tag{2.165}$$

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

$$\mathbf{E}_{\mathrm{t}}^2\left(t\right)=\frac{\left|\mathbf{E}\right|^2}{2}\left[1+\left|\eta\right|\cos\left(2\omega t-\vartheta\right)\right]\ . \tag{2.166}$$

Thus, while $|\eta|$ determines the eccentricity

$$e=\sqrt{1-\frac{(\mathbf{E}_\mathrm{t}^2)_{\min}}{(\mathbf{E}_\mathrm{t}^2)_{\max}}}=\sqrt{\frac{2|\eta|}{1+|\eta|}}\;, \tag{2.167}$$

the phase ϑ of η determines the phase of time oscillations.

Exercise 2.11.2. Show that

$$\left(\mathbf{E}_\mathrm{t}^2\right)_{\min,\max}=\frac{1}{2}\left(\mathbf{a}^2+\mathbf{b}^2\right)\pm\frac{1}{2}\sqrt{4\left(\mathbf{a}\cdot\mathbf{b}\right)^2+\left(\mathbf{a}^2-\mathbf{b}^2\right)^2}\;. \tag{2.168}$$

Solution 2.11.2. In terms of $\mathbf{a}$ and $\mathbf{b}$ one has [see Eq. (2.154)]

$$\mathbf{E}_\mathrm{t}^2\left(t\right)=\mathbf{a}^2\cos^2\left(\omega t\right)+\mathbf{b}^2\sin^2\left(\omega t\right)+\mathbf{a}\cdot\mathbf{b}\sin\left(2\omega t\right)\;. \tag{2.169}$$

The extremum points of $\mathbf{E}_\mathrm{t}^2$ are found by solving

$$0=\frac{\mathrm{d}\mathbf{E}_\mathrm{t}^2}{\mathrm{d}\left(\omega t\right)}=-\mathbf{a}^2\sin\left(2\omega t\right)+\mathbf{b}^2\sin\left(2\omega t\right)+2\mathbf{a}\cdot\mathbf{b}\cos\left(2\omega t\right)\;, \tag{2.170}$$

thus

$$\tan\left(2\omega t\right)=\frac{2\mathbf{a}\cdot\mathbf{b}}{\mathbf{a}^2-\mathbf{b}^2}\;. \tag{2.171}$$

Rewriting Eq. (2.169) as

$$\mathbf{E}_\mathrm{t}^2\left(t\right)=\frac{1}{2}\left(\mathbf{a}^2+\mathbf{b}^2\right)+\frac{1}{2}\cos\left(2\omega t\right)\left(\mathbf{a}^2-\mathbf{b}^2\right)+\mathbf{a}\cdot\mathbf{b}\sin\left(2\omega t\right)\;, \tag{2.172}$$

and using Eq. (2.171) together with the identities

$$\sin x=\pm\frac{\tan x}{\sqrt{\tan^2 x+1}}\;, \tag{2.173}$$

and

$$\cos x=\pm\frac{1}{\sqrt{\tan^2 x+1}}\;, \tag{2.174}$$

lead to

$$\begin{aligned}\left(\mathbf{E}_\mathrm{t}^2\right)_{\min,\max}&=\frac{1}{2}\left(\mathbf{a}^2+\mathbf{b}^2\right)\pm\frac{1}{2}\frac{1}{\sqrt{\left(\frac{2\mathbf{a}\cdot\mathbf{b}}{\mathbf{a}^2-\mathbf{b}^2}\right)^2+1}}\left[\left(\mathbf{a}^2-\mathbf{b}^2\right)+\mathbf{a}\cdot\mathbf{b}\frac{4\mathbf{a}\cdot\mathbf{b}}{\mathbf{a}^2-\mathbf{b}^2}\right]\\&=\frac{1}{2}\left(\mathbf{a}^2+\mathbf{b}^2\right)\pm\frac{1}{2}\sqrt{4\left(\mathbf{a}\cdot\mathbf{b}\right)^2+\left(\mathbf{a}^2-\mathbf{b}^2\right)^2}\;.\end{aligned} \tag{2.175}$$

2.12 Problems

1. Calculate the electric field $\mathbf{E}$, magnetic field $\mathbf{B}$ and the Poynting vector $\mathbf{S}$ generated by a point particle of charge q moving in the x direction at a fixed velocity u.
2. Consider an infinite wire carrying charge per unit length λ, which flows along the wire with a constant velocity u. Let E and B be the magnitude of the electric and magnetic fields, respectively, at a distance l from the wire. Calculate E and B using two methods. The first one is based on the integral representation of the Maxwell's equations given by Eqs. (2.69), (2.70), (2.71) and (2.72). The second method is based on Eqs. (2.193) and (2.194) (these equations were derived for the problem of a point charge moving at a constant velocity u). Compare the results obtained from these two methods.
3. Consider an uncharged dielectric sphere having a homogeneous permittivity ϵ and radius R. Calculate the electric field $\mathbf{E}$ inside and outside the sphere given that far from the sphere $\mathbf{E} = E_0\hat{\mathbf{z}}$, where E_0 is a constant.
4. Consider a sphere of radius R in vacuum having uniform permanent magnetization given by $\mathbf{M} = M_0\hat{\mathbf{z}}$, where M_0 is a constant and $\hat{\mathbf{z}}$ is a unit vector. Calculate the magnetic fields $\mathbf{B}$ and $\mathbf{H}$.
5. **The Drude model** - Consider a conductor containing charge carriers having charge q and mass m in 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)]
$$\frac{\mathrm{d}\mathbf{p}}{\mathrm{d}t} + \frac{\mathbf{p}}{\tau_{\mathrm{tr}}} = q\mathbf{E} , \qquad (2.176)$$
where $\mathbf{p}$ is the momentum per electron and where τ_{tr} is the so-called scattering time. Calculate the frequency dependent effective dielectric coefficient $\epsilon_{\mathrm{eff}}(\omega)$ of the conductor.
6. **Fresnel equations** - Consider a plane wave of wave vector $\mathbf{k}_{\mathrm{i}}$ striking the planar interface between two lossless materials having refractive indices $n_l = \sqrt{\epsilon_l \mu_l}$, where for the material hosting the incident wave $l = 1$ and for the other material $l = 2$. The plane containing $\mathbf{k}_{\mathrm{i}}$ and $\hat{\mathbf{n}}$, where $\hat{\mathbf{n}}$ is a unit vector normal to the interface, is called the plane of incidence. Let θ_{i} be the angle between $\mathbf{k}_{\mathrm{i}}$ and $\hat{\mathbf{n}}$. 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).
7. Consider a layer of width d and a refractive index $n_3 = \sqrt{\epsilon_3 \mu_3}$ sandwiched between two semi-infinite media having refractive index $n_1 = \sqrt{\epsilon_1 \mu_1}$

and $n_2 = \sqrt{\epsilon_2 \mu_2}$, 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 ω corresponding to surface modes having s-polarization and an equation for the case of p-polarization.

8. **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 ω 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.
   b) Calculate the Casimir force for the case where both layers 1 and 2 have metallic dielectric coefficient given by [see Eq. (2.229)]
   $$\epsilon_{1,2}(\omega) = 1 - \frac{\omega_{\mathrm{p}}^2}{\omega^2} , \tag{2.177}$$
   where ω_{p} is the plasma frequency, $\epsilon_3 = 1$ and $\mu_1 = \mu_2 = \mu_3 = 1$.
9. **Stokes parameters** - Consider a monochromatic electromagnetic plane wave propagating in vacuum along the z axis. The components of the electric field vector $\mathbf{E} = (E_x, E_y, E_z)$ are assumed to be given by
   $$E_x = E_{x0} \cos\left(\omega \frac{z - ct}{c} + \delta_x\right) , \tag{2.178}$$
   $$E_y = E_{y0} \cos\left(\omega \frac{z - ct}{c} + \delta_y\right) , \tag{2.179}$$
   $$E_z = 0 , \tag{2.180}$$
   where ω, E_{x0}, E_{y0}, δ_x and δ_y are constants. The so-called Stokes parameters are defined by
   $$S_0 = E_{x0}^2 + E_{y0}^2 , \tag{2.181}$$
   $$S_1 = E_{x0}^2 - E_{y0}^2 , \tag{2.182}$$
   $$S_2 = 2E_{x0}E_{y0} \cos(\delta_x - \delta_y) , \tag{2.183}$$
   $$S_3 = 2E_{x0}E_{y0} \sin(\delta_x - \delta_y) . \tag{2.184}$$
   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 $c\boldsymbol{\beta}$, where the vector $\boldsymbol{\beta}$ is expressed as $\boldsymbol{\beta} = \beta\hat{\boldsymbol{\beta}}$, where $\beta = |\boldsymbol{\beta}|$, and the unit vector $\hat{\boldsymbol{\beta}} = (\sin\theta, 0, \cos\theta)$ is assumed to lie in the xz plane.
10. **The Drude-Lorentz model** - Consider light propagating in the z direction in a medium containing resonators with number density N (resonators per unit volume). Each resonator has mass m, charge e, damping

rate γ and angular frequency ω_0. A magnetic field given by $\mathbf{H} = H_0\hat{\mathbf{z}}$ is externally applied in the direction of propagation.

a) Calculate the indices of refraction n_+ and n_- corresponding to clockwise and counter clockwise circular states of polarization, respectively. Assume that $\gamma \ll |\omega - \omega_0|$, where ω is the angular frequency of the propagating light, and $eH_0/(mc) \ll \omega_0$.
b) Calculate the Verdet constant V, which is defined by $V = \Delta_\phi/(H_0 z)$, where Δ_ϕ is the rotation angle of linear polarization traveling a distance z. This polarization rotation is known as the Faraday effect.
c) Calculate the magnetization $\mathbf{M}$ generated by the motion of the resonators. This optically-induced magnetization is known as the inverse Faraday effect.

11. Show that the scalars $\mathbf{E}\cdot\mathbf{B}$, $\mathbf{E}\cdot\mathbf{E}-\mathbf{B}\cdot\mathbf{B}$, $\mathbf{D}\cdot\mathbf{H}$, $\mathbf{D}\cdot\mathbf{D}-\mathbf{H}\cdot\mathbf{H}$ and $\mathbf{E}\cdot\mathbf{D}-\mathbf{B}\cdot\mathbf{H}$ are all Lorentz invariant.
12. The magnetic field $\mathbf{B}(t,\mathbf{x})$ vanishes for any position $\mathbf{x}$ and at any time t in a given inertial frame S and the electric field $\mathbf{E}'(t',\mathbf{x}')$ vanishes for any position $\mathbf{x}'$ and at any t' in another inertial frame S', which moves at a constant velocity with respect to S. What can be said about the electric field $\mathbf{E}(t,\mathbf{x})$ in the inertial frame S?
13. **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

$$A(t,\mathbf{x}) = \int \mathrm{d}^3\mathbf{x}'\, \frac{J\left(t - \frac{|\mathbf{x}-\mathbf{x}'|}{c}, \mathbf{x}'\right)}{c\,|\mathbf{x}-\mathbf{x}'|}\ . \tag{2.185}$$

14. Consider a point particle having charge q. The location $\mathbf{x}'(t')$ of the particle at time t' in Cartesian coordinates is given by $\mathbf{x}'(t') = r_0(\cos\omega t', \sin\omega t', 0)$, where both $r_0 > 0$ and $\omega > 0$ are constants. Calculate the The Leinard–Wiechert potential $A(t,\mathbf{x})$ (2.185) at the point $\mathbf{x} = (0,0,z)$.
15. Calculate the Leinard–Wiechert potential for a point particle having charge q, which moves along the x direction at a fixed velocity u. Use the result to calculate the electric $\mathbf{E}$ and magnetic $\mathbf{B}$ fields generated by the moving particle.
16. **Far field** - Consider charge distribution having density $\rho(t,\mathbf{x}')$. The electric dipole moment $\mathbf{p}$ is given by

$$\mathbf{p} = \int \mathrm{d}^3\mathbf{x}'\ \mathbf{x}'\rho(t,\mathbf{x}')\ . \tag{2.186}$$

The charge distribution is localized inside a sphere of radius r_c centered at the origin of spatial coordinates (i.e. $|\mathbf{x}'| < r_\mathrm{c}$). Let ω_c be a characteristic angular frequency of radiation emitted from the distribution due to motion of charges. Express the electric $\mathbf{E}$ and magnetic $\mathbf{B}$ fields at the space-time point $(t,\mathbf{x})$ in terms of $\mathbf{p}$ in the so-called far field limit, for which $|\mathbf{x}| \gg r_\mathrm{c}$ (dipole approximation) and $|\mathbf{x}| \gg c/\omega_\mathrm{c}$, and calculate the total emitted radiative power P.

2.13 Solutions

1. In a frame commoving with the particle the electric field $\mathbf{E}$ is given by the Coulomb's law

$$\mathbf{E}' = \frac{q\mathbf{r}'}{|\mathbf{r}'|^3} , \tag{2.187}$$

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

$$\begin{aligned}\mathbf{E} &= \mathbf{E}'_{\parallel} + \gamma \left(\mathbf{E}'_{\perp} - \boldsymbol{\beta} \times \mathbf{B}'_{\perp}\right) \\ &= \frac{q x'_1 \hat{\mathbf{x}}_1}{\left(x'^2_1 + x'^2_2 + x'^2_3\right)^{3/2}} + \gamma \frac{q \left(x'_2 \hat{\mathbf{x}}_2 + x'_3 \hat{\mathbf{x}}_3\right)}{\left(x'^2_1 + x'^2_2 + x'^2_3\right)^{3/2}} ,\end{aligned} \tag{2.188}$$

and

$$\begin{aligned}\mathbf{B} &= \mathbf{B}'_{\parallel} + \gamma \left(\mathbf{B}'_{\perp} + \boldsymbol{\beta} \times \mathbf{E}'_{\perp}\right) \\ &= \gamma \boldsymbol{\beta} \times \frac{q \left(x'_2 \hat{\mathbf{x}}_2 + x'_3 \hat{\mathbf{x}}_3\right)}{\left(x'^2_1 + x'^2_2 + x'^2_3\right)^{3/2}} ,\end{aligned} \tag{2.189}$$

where

$$\boldsymbol{\beta} = \frac{u}{c} \hat{\mathbf{x}}_1 , \tag{2.190}$$

and where

$$\gamma = \frac{1}{\sqrt{1 - \left(\frac{u}{c}\right)^2}} . \tag{2.191}$$

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

$$\begin{pmatrix} ct' \\ x'_1 \end{pmatrix} = \gamma \begin{pmatrix} 1 & -\frac{u}{c} \\ -\frac{u}{c} & 1 \end{pmatrix} \begin{pmatrix} ct \\ x_1 \end{pmatrix} , \tag{2.192}$$

and therefore

$$\mathbf{E} = \frac{q\gamma}{r_0^3} \mathbf{x}_0 , \tag{2.193}$$

and

$$\mathbf{B} = \frac{q\gamma}{r_0^3} \boldsymbol{\beta} \times \mathbf{x}_0 , \tag{2.194}$$

where

$$\mathbf{x}_0 = \mathbf{x} - \mathbf{u}t \,, \tag{2.195}$$

and where

$$r_0 = \sqrt{\gamma^2 (x_1 - ut)^2 + x_2^2 + x_3^2} \,. \tag{2.196}$$

With the help of Eqs. (2.93), (2.193), (2.194) and (3.65) one finds that the Poynting vector is given by

$$\mathbf{S} = \frac{cq^2\gamma^2}{4\pi} \frac{(\mathbf{x}_0 \cdot \mathbf{x}_0)\, \boldsymbol{\beta} - (\mathbf{x}_0 \cdot \boldsymbol{\beta})\, \mathbf{x}_0}{r_0^6} \,, \tag{2.197}$$

or

$$\mathbf{S} = \frac{q^2\gamma^2 u}{4\pi} \frac{\left(x_2^2 + x_3^2\right) \hat{\mathbf{x}}_1 - (x_1 - ut)\left(x_2 \hat{\mathbf{x}}_2 + x_3 \hat{\mathbf{x}}_3\right)}{\left(\gamma^2 (x_1 - ut)^2 + x_2^2 + x_3^2\right)^3} \,. \tag{2.198}$$

Note that no radiation is emitted from the moving particle. This can be seen from the energy conservation law (2.92), which for this case (the conductivity σ vanishes) reads

$$\int_S \mathbf{S} \cdot \mathrm{d}\mathbf{s} + \frac{\partial}{\partial t} \int_V u \, \mathrm{d}v = 0 \,, \tag{2.199}$$

and from the fact that $|\mathbf{S}|$ roughly decays as $|\mathbf{x}|^{-4}$ [see Eq. (2.198)].

2. In the first method, using Eq. (2.71) one finds that $E = 2\lambda/l$ [see Eqs. (2.17) and (2.68)], and using Eq. (2.69) one finds that $B = (u/c)(2\lambda/l)$ [see Eqs. (2.18) and (2.67)]. In the second method, using Eqs. (2.193) and (2.194) one finds that $E = \lambda\gamma u l I$ and $B = (u/c) E$, where the integral I is given by [the charge q in Eqs. (2.193) and (2.194) is replaced by $\lambda u \times \mathrm{d}t$, and integration over time t is performed]

$$I = \int_{-\infty}^{\infty} \frac{\mathrm{d}t}{\left((\gamma u t)^2 + l^2\right)^{3/2}} \,, \tag{2.200}$$

thus using the definite integral

$$\int_{-\infty}^{\infty} \frac{\mathrm{d}q}{(1+q^2)^{3/2}} = 2 \,, \tag{2.201}$$

one finds that $E = 2\lambda/l$ and $B = (u/c)(2\lambda/l)$, in agreement with the first method.

3. The coordinates are chosen such that the center of the sphere is located at the origin. Since the magnetic induction $\mathbf{B}$ vanishes for this problem of electrostatics $\mathbf{E}$ can be expressed in terms of a scalar potential ϕ as [see Eq. (2.13)]

$$\mathbf{E} = -\boldsymbol{\nabla}\phi \,. \tag{2.202}$$

The transformation from Cartesian to spherical coordinates is given by

$$x = r\sin\theta\cos\varphi \,, \tag{2.203}$$

$$y = r\sin\theta\sin\varphi \,, \tag{2.204}$$

$$z = r\cos\theta \,, \tag{2.205}$$

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

$$\begin{aligned}\boldsymbol{\nabla} &= \hat{\mathbf{x}}\frac{\partial}{\partial x} + \hat{\mathbf{y}}\frac{\partial}{\partial y} + \hat{\mathbf{z}}\frac{\partial}{\partial z} \\ &= \hat{\mathbf{r}}\frac{\partial}{\partial r} + \hat{\boldsymbol{\theta}}\frac{1}{r}\frac{\partial}{\partial \theta} + \hat{\boldsymbol{\varphi}}\frac{1}{r\sin\theta}\frac{\partial}{\partial \varphi} \,.\end{aligned} \tag{2.206}$$

By symmetry, the scalar potential ϕ is expected to be independent on φ, and thus it can be expressed as

$$\phi(r,\theta) = \begin{cases} \phi_{\mathrm{in}}(r,\theta) & r < R \\ \phi_{\mathrm{out}}(r,\theta) & r > R \end{cases} \,. \tag{2.207}$$

The requirement that $\mathbf{E} = E_0\hat{\mathbf{z}}$ far from the sphere implies that $\phi_{\mathrm{out}} \simeq -E_0 z = -E_0 r\cos\theta$ when $r \gg R$. On the surface of the sphere the boundary condition (2.78) reads

$$\hat{\mathbf{n}} \times \boldsymbol{\nabla}\phi_{\mathrm{out}} = \hat{\mathbf{n}} \times \boldsymbol{\nabla}\phi_{\mathrm{in}} \,, \tag{2.208}$$

and the boundary condition (2.79) reads [recall Eq. (2.85), which reads $\mathbf{D} = \epsilon\mathbf{E}$]

$$\hat{\mathbf{n}} \cdot \boldsymbol{\nabla}\phi_{\mathrm{out}} = \epsilon\hat{\mathbf{n}} \cdot \boldsymbol{\nabla}\phi_{\mathrm{in}} \,, \tag{2.209}$$

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

$$\frac{\partial\phi_{\mathrm{out}}}{\partial\theta} = \frac{\partial\phi_{\mathrm{in}}}{\partial\theta} \,, \tag{2.210}$$

and

$$\frac{\partial\phi_{\mathrm{out}}}{\partial r} = \epsilon\frac{\partial\phi_{\mathrm{in}}}{\partial r} \,. \tag{2.211}$$

All these requirements are satisfied by

$$\frac{\phi}{E_0} = \begin{cases} -\frac{3}{2+\epsilon} r\cos\theta & r < R \\ \left(\frac{\epsilon-1}{\epsilon+2}\frac{R^3}{r^2} - r\right)\cos\theta & r > R \end{cases} \,, \tag{2.212}$$

and thus [see Eqs. (2.202) and (2.206)]

$$\frac{\mathbf{E}}{E_0} = \begin{cases} 3\frac{\hat{\mathbf{r}}\cos\theta - \hat{\boldsymbol{\theta}}\sin\theta}{2+\epsilon} & r < R \\ \hat{\mathbf{r}}\left(1 + \frac{2(\epsilon-1)}{\frac{r^3}{R^3}(2+\epsilon)}\right)\cos\theta + \hat{\boldsymbol{\theta}}\left(\frac{\epsilon-1}{\frac{r^3}{R^3}(2+\epsilon)} - 1\right)\sin\theta & r > R \end{cases} . \quad (2.213)$$

4. For this magnetostatic problem the filed $\mathbf{H}$ can be expressed in terms of a scalar function φ_{m} as $\mathbf{H} = -\boldsymbol{\nabla}\varphi_{\mathrm{m}}$, since $\mathbf{J}_{\mathrm{total}} = \mathbf{J}_{\mathrm{mag}}$ [see Eq. (2.8)], and thus $\boldsymbol{\nabla} \times \mathbf{H} = 0$ [see Eqs. (2.12)]. In addition, the following holds $\boldsymbol{\nabla} \cdot \mathbf{B} = 0$ [see Eq. (2.15)] and $\mathbf{H} = \mathbf{B} - 4\pi\mathbf{M}$ [see Eq. (2.18)], and thus φ_{m} satisfies the following Poisson equation [see Eq. (1.5)]

$$\boldsymbol{\nabla}^2\varphi_{\mathrm{m}} = -4\pi\rho_{\mathrm{m}} , \quad (2.214)$$

where $\rho_{\mathrm{m}} = -\boldsymbol{\nabla} \cdot \mathbf{M}$. The solution is given by [see Eq. (1.3)]

$$\varphi_{\mathrm{m}}(\mathbf{r}) = \int \mathrm{d}^3\mathbf{r}' \frac{\rho_{\mathrm{m}}(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} . \quad (2.215)$$

On the surface of the sphere the discontinuity of $\mathbf{M}$ gives rise to an effective surface charge density σ_{m} given in spherical coordinates [see Eqs. (2.68), (2.203), (2.204) and (2.205)] by $\sigma_{\mathrm{m}} = M_0\cos\theta$ (the sphere's center is assumed to be located at the origin). Using the so-called addition theorem, which is given by

$$\frac{1}{|\mathbf{r} - \mathbf{r}'|} = \sum_{l=0}^{\infty} \frac{4\pi}{2l+1} \frac{r_<^l}{r_>^{l+1}} \sum_{m=-l}^{l} Y_l^{m*}(\theta', \varphi') Y_l^m(\theta, \varphi) , \quad (2.216)$$

where in spherical coordinates $\mathbf{r} = r(\sin\theta\cos\varphi, \sin\theta\sin\varphi, \cos\theta)$, $\mathbf{r}' = r'(\sin\theta'\cos\varphi', \sin\theta'\sin\varphi', \cos\theta')$, $r_< = \min(r, r')$, $r_> = \max(r, r')$, and $Y_l^m(\theta, \varphi)$ are the spherical harmonics functions, together with the orthogonality relation

$$\int_0^{2\pi} \mathrm{d}\varphi' \int_{-1}^{1} \mathrm{d}(\cos\theta') Y_{l'}^{m'*}(\theta', \varphi') Y_l^m(\theta', \varphi') = \delta_{l.l'}\delta_{m.m'} , \quad (2.217)$$

one finds that (note that $\cos\theta' = Y_1^0(\theta', \phi') / \sqrt{3/4\pi}$)

$$\varphi_{\mathrm{m}}(\mathbf{r}) = \begin{cases} \frac{4\pi M_0 r\cos\theta}{3} = \frac{4\pi M_0 z}{3} & r < R \\ \frac{4\pi M_0 R^3\cos\theta}{3r^2} & r \geq R \end{cases} . \quad (2.218)$$

Using the relation $\mathbf{H} = -\boldsymbol{\nabla}\varphi_{\mathrm{m}}$ one obtains [see Eq. (2.206)]

$$\mathbf{H}(\mathbf{r}) = \begin{cases} -\frac{4\pi M_0}{3}\hat{\mathbf{z}} & r < R \\ \frac{4\pi R^3 M_0}{3r^3}\left(2\hat{\mathbf{r}}\cos\theta + \hat{\boldsymbol{\theta}}\sin\theta\right) & r \geq R \end{cases} , \quad (2.219)$$

or (note that $\hat{\mathbf{r}}\cos\theta - \hat{\boldsymbol{\theta}}\sin\theta = \hat{\mathbf{z}}$, where $\hat{\mathbf{r}}$ and $\hat{\boldsymbol{\theta}}$ are unit vectors in the radial and azimuthal directions, respectively)

$$\mathbf{H}(\mathbf{r}) = \begin{cases} -\frac{\mathbf{m}}{R^3} & r < R \\ -\frac{\mathbf{m}-3\hat{\mathbf{r}}(\hat{\mathbf{r}}\cdot\mathbf{m})}{r^3} & r \geq R \end{cases} , \tag{2.220}$$

where the dipole moment $\mathbf{m}$ is given by

$$\mathbf{m} = \frac{4\pi R^3}{3}\mathbf{M} , \tag{2.221}$$

and thus

$$\mathbf{B}(\mathbf{r}) = \begin{cases} \frac{2\mathbf{m}}{R^3} & r < R \\ -\frac{\mathbf{m}-3\hat{\mathbf{r}}(\hat{\mathbf{r}}\cdot\mathbf{m})}{r^3} & r \geq R \end{cases} . \tag{2.222}$$

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

$$\mathbf{p} = \frac{m}{qn_{\mathrm{cc}}}\mathbf{J} , \tag{2.223}$$

Eq. (2.176) yields

$$\frac{m}{q^2 n_{\mathrm{cc}}}\left(\frac{\partial \mathbf{J}}{\partial t} + \frac{1}{\tau_{\mathrm{tr}}}\mathbf{J}\right) = \mathbf{E} . \tag{2.224}$$

When harmonic time dependency at angular frequency ω is assumed Eq. (2.224) yields

$$\frac{m}{q^2 n_{\mathrm{cc}}}\left(-i\omega + \frac{1}{\tau_{\mathrm{tr}}}\right)\mathbf{J}(\omega) = \mathbf{E}(\omega) , \tag{2.225}$$

and thus the conductivity σ is given by [see Eq. (2.91)]

$$\sigma(\omega) = \frac{\sigma_0}{1 - i\omega\tau_{\mathrm{tr}}} , \tag{2.226}$$

where

$$\sigma_0 = \frac{q^2 n_{\mathrm{cc}} \tau_{\mathrm{tr}}}{m} , \tag{2.227}$$

and therefore the effective dielectric coefficient ϵ_{eff} is given by [see Eq. (2.132)]

$$\epsilon_{\mathrm{eff}}(\omega) = 1 + \frac{4\pi i}{\omega}\frac{\sigma_0}{1 - i\omega\tau_{\mathrm{tr}}} . \tag{2.228}$$

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

$$\epsilon_{\mathrm{eff}}(\omega) = 1 - \frac{\omega_{\mathrm{p}}^2}{\omega^2} , \tag{2.229}$$

where ω_{p}, which is given by

$$\omega_{\mathrm{p}}^2 = \frac{4\pi q^2 n_{\mathrm{cc}}}{m} , \tag{2.230}$$

is the so-called plasma frequency.

6. 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 $\hat{\mathbf{n}} = \hat{\mathbf{z}}$ and $\hat{\mathbf{x}}$. Consider a solution for the electric field $\mathbf{E}$ composed of incident, reflected and refracted plane waves having wave vectors $\mathbf{k}_{\mathrm{i}}$, $\mathbf{k}_{\mathrm{r}}$ and $\mathbf{k}_{\mathrm{t}}$, respectively. For the case of s-polarization the solution is expressed as [see Eq. (2.151)]

$$\mathbf{E}_{\mathrm{s}} = \begin{cases} E_{\mathrm{i}} \left(e^{i\mathbf{k}_{\mathrm{i}}\cdot\mathbf{r}} + r_{\mathrm{s}} e^{i\mathbf{k}_{\mathrm{r}}\cdot\mathbf{r}} \right) \hat{\mathbf{y}} & z < 0 \\ t_{\mathrm{s}} E_{\mathrm{i}} e^{i\mathbf{k}_{\mathrm{t}}\cdot\mathbf{r}} \hat{\mathbf{y}} & z > 0 \end{cases} , \tag{2.231}$$

and for the case of p-polarization as

$$\mathbf{E}_{\mathrm{p}} = \begin{cases} E_{\mathrm{i}} \left(e^{i\mathbf{k}_{\mathrm{i}}\cdot\mathbf{r}} \frac{\mathbf{k}_{\mathrm{i}}\times\hat{\mathbf{y}}}{k_{\mathrm{i}}} + r_{\mathrm{p}} e^{i\mathbf{k}_{\mathrm{r}}\cdot\mathbf{r}} \frac{\mathbf{k}_{\mathrm{r}}\times\hat{\mathbf{y}}}{k_{\mathrm{r}}} \right) & z < 0 \\ E_{\mathrm{i}} t_{\mathrm{p}} e^{i\mathbf{k}_{\mathrm{t}}\cdot\mathbf{r}} \frac{\mathbf{k}_{\mathrm{t}}\times\hat{\mathbf{y}}}{k_{\mathrm{t}}} & z > 0 \end{cases} , \tag{2.232}$$

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-polarization) and E_{i} is the amplitude of incident wave. The boundary condition (2.78) can be satisfied for every point in the plane $z = 0$ only when $\mathbf{k}_{\mathrm{i}}$, $\mathbf{k}_{\mathrm{r}}$ and $\mathbf{k}_{\mathrm{t}}$ have the same tangential component. This requirement is satisfied by expressing $\mathbf{k}_{\mathrm{i}}$, $\mathbf{k}_{\mathrm{r}}$ and $\mathbf{k}_{\mathrm{t}}$ in terms of the corresponding angles θ_{i}, θ_{r} and θ_{t} as [see Eq. (2.151)]

$$\mathbf{k}_{\mathrm{i}} = \frac{n_1\omega}{c} \left(\sin\theta_{\mathrm{i}}, 0, \cos\theta_{\mathrm{i}}\right) , \tag{2.233}$$

$$\mathbf{k}_{\mathrm{r}} = \frac{n_1\omega}{c} \left(\sin\theta_{\mathrm{r}}, 0, -\cos\theta_{\mathrm{r}}\right) , \tag{2.234}$$

$$\mathbf{k}_{\mathrm{t}} = \frac{n_2\omega}{c} \left(\sin\theta_{\mathrm{t}}, 0, \cos\theta_{\mathrm{t}}\right) , \tag{2.235}$$

where θ_{r} is related to θ_{i} by the so-called law of reflection

$$\theta_{\mathrm{r}} = \theta_{\mathrm{i}} , \tag{2.236}$$

and θ_{t} is related to θ_{i} by the so-called Snell's law

$$n_1 \sin\theta_{\mathrm{i}} = n_2 \sin\theta_{\mathrm{t}} . \tag{2.237}$$

The magnetic field $\mathbf{H}$ is related to $\mathbf{E}$ by Eq. (2.139), which reads

$$\mathbf{H} = -i\frac{c}{\omega\mu}\boldsymbol{\nabla}\times\mathbf{E}\ , \tag{2.238}$$

thus [see Eqs. (1.96), (2.150), (2.231) and (2.232)]

$$\mathbf{H}_{\mathrm{s}} = \begin{cases} \frac{E_{\mathrm{i}}n_1}{\mu_1}\left(e^{i\mathbf{k}_{\mathrm{i}}\cdot\mathbf{r}}\frac{\mathbf{k}_{\mathrm{i}}\times\hat{\mathbf{y}}}{k_{\mathrm{i}}} + r_{\mathrm{s}}e^{i\mathbf{k}_{\mathrm{r}}\cdot\mathbf{r}}\frac{\mathbf{k}_{\mathrm{r}}\times\hat{\mathbf{y}}}{k_{\mathrm{r}}}\right) & z<0 \\ \frac{E_{\mathrm{i}}n_2}{\mu_2}t_{\mathrm{s}}e^{i\mathbf{k}_{\mathrm{t}}\cdot\mathbf{r}}\frac{\mathbf{k}_{\mathrm{t}}\times\hat{\mathbf{y}}}{k_{\mathrm{t}}} & z>0 \end{cases}\ , \tag{2.239}$$

and

$$\mathbf{H}_{\mathrm{p}} = \begin{cases} -\frac{E_{\mathrm{i}}n_1}{\mu_1}\left(e^{i\mathbf{k}_{\mathrm{i}}\cdot\mathbf{r}} + r_{\mathrm{p}}e^{i\mathbf{k}_{\mathrm{r}}\cdot\mathbf{r}}\right)\hat{\mathbf{y}} & z<0 \\ -\frac{E_{\mathrm{i}}n_2}{\mu_2}t_{\mathrm{p}}e^{i\mathbf{k}_{\mathrm{t}}\cdot\mathbf{r}}\hat{\mathbf{y}} & z>0 \end{cases}\ . \tag{2.240}$$

The boundary conditions (2.77) and (2.78) yield for the case of s-polarization [see Eqs. (2.231) and (2.239)]

$$\sqrt{\frac{\epsilon_1}{\mu_1}}(1-r_{\mathrm{s}})\cos\theta_{\mathrm{r}} = \sqrt{\frac{\epsilon_2}{\mu_2}}t_{\mathrm{s}}\cos\theta_{\mathrm{t}}\ , \tag{2.241}$$

$$1+r_{\mathrm{s}} = t_{\mathrm{s}}\ , \tag{2.242}$$

and for the case of p-polarization [see Eqs. (2.232) and (2.240)]

$$\sqrt{\frac{\epsilon_1}{\mu_1}}(1+r_{\mathrm{p}}) = \sqrt{\frac{\epsilon_2}{\mu_2}}t_{\mathrm{p}}\ , \tag{2.243}$$

$$(1-r_{\mathrm{p}})\cos\theta_{\mathrm{r}} = t_{\mathrm{p}}\cos\theta_{\mathrm{t}}\ . \tag{2.244}$$

The solutions are given by

$$r_{\mathrm{s}} = \frac{\sqrt{\frac{\epsilon_1}{\mu_1}}\cos\theta_{\mathrm{i}} - \sqrt{\frac{\epsilon_2}{\mu_2}}\cos\theta_{\mathrm{t}}}{\sqrt{\frac{\epsilon_1}{\mu_1}}\cos\theta_{\mathrm{i}} + \sqrt{\frac{\epsilon_2}{\mu_2}}\cos\theta_{\mathrm{t}}}\ , \tag{2.245}$$

and

$$r_{\mathrm{p}} = \frac{\sqrt{\frac{\epsilon_2}{\mu_2}}\cos\theta_{\mathrm{i}} - \sqrt{\frac{\epsilon_1}{\mu_1}}\cos\theta_{\mathrm{t}}}{\sqrt{\frac{\epsilon_2}{\mu_2}}\cos\theta_{\mathrm{i}} + \sqrt{\frac{\epsilon_1}{\mu_1}}\cos\theta_{\mathrm{t}}}\ . \tag{2.246}$$

Note that when total internal reflection occurs, i.e. when $n_1 > n_2$ and [see Eq. (2.237)]

$$\sin\theta_{\mathrm{i}} > \frac{n_2}{n_1}\ , \tag{2.247}$$

one has $|r_{\mathrm{s}}|^2 = |r_{\mathrm{p}}|^2 = 1$ [note that Eq. (2.237) yields no real solution for θ_{t} for this case]. Consider the case where $\mu_1 = \mu_2$. For that case $|r_{\mathrm{p}}|^2 = 0$ when $\theta_{\mathrm{i}} = \theta_{\mathrm{B}}$, where θ_{B} is the so-called Brewster's angle, which is given by [see Eqs. (2.237) and (2.246)]

$$\theta_{\mathrm{B}} = \tan^{-1}\frac{n_2}{n_1}\ . \tag{2.248}$$

7. 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. (2.231)]

$$\mathbf{E} = \hat{\mathbf{y}} e_y(z) e^{i\boldsymbol{\kappa}\cdot\boldsymbol{\rho}} , \tag{2.249}$$

where $\boldsymbol{\kappa} = (k_x, k_y)$ and $\boldsymbol{\rho} = (x, y)$. The Helmholtz equation (2.151) implies that

$$\frac{e_y''}{e_y} = K^2 , \tag{2.250}$$

where

$$K^2 = \kappa^2 - \left(\frac{n\omega}{c}\right)^2 . \tag{2.251}$$

Note that for the case of surface modes $K^2 > 0$. The magnetic field $\mathbf{H}$ is given by [see Eqs. (2.139) and (2.150)]

$$\begin{aligned}\mathbf{H} &= \frac{1}{ik_0\mu}\boldsymbol{\nabla}\times\mathbf{E} \\ &= \frac{\left(-e_y', 0, 0\right) + i\left(0, 0, k_x e_y\right)}{ik_0\mu} e^{i\boldsymbol{\kappa}\cdot\boldsymbol{\rho}} ,\end{aligned} \tag{2.252}$$

where $k_0 = \omega/c$ [see Eq. (2.142)]. Next, consider an interface between two materials at a plane of constant z. The Snell's law (2.237) implies that the lateral wave vector $\boldsymbol{\kappa}$ obtains the same value on both sides of the interface. Thus, for this case of s-polarization the boundary condition (2.78) implies that e_y is continuous, and the boundary condition (2.77) implies that $\mu^{-1}e_y'$ is continuous. For the trilayer, the refractive index n is taken to be given by

$$n(z) = \begin{cases} n_1 = \sqrt{\epsilon_1\mu_1} & z < 0 \\ n_3 = \sqrt{\epsilon_3\mu_3} & 0 \le z \le d \\ n_2 = \sqrt{\epsilon_2\mu_2} & z < d \end{cases} . \tag{2.253}$$

Consider a solution having the form [see Eq. (2.250)]

$$e_y(z) = \begin{cases} Ae^{K_1 z} & z < 0 \\ Be^{K_3 z} + Ce^{-K_3 z} & 0 \le z \le d \\ De^{-K_2 z} & z < d \end{cases} , \tag{2.254}$$

where [see Eq. (2.251)]

$$K_l = \sqrt{\kappa^2 - \left(\frac{n_l\omega}{c}\right)^2} . \tag{2.255}$$

In a matrix form the boundary conditions at the interfaces $z=0$ and $z=d$ can be expressed as

$$M_{\mathrm{s}}\begin{pmatrix} A & B & C & D \end{pmatrix}^{\mathrm{T}}=0\ , \tag{2.256}$$

where

$$M_{\mathrm{s}}=\begin{pmatrix} 1 & -1 & -1 & 0 \\ 0 & e^{K_3 d} & e^{-K_3 d} & -e^{-K_3 d} \\ \frac{K_1}{\mu_1} & -\frac{K_3}{\mu_3} & \frac{K_3}{\mu_3} & 0 \\ 0 & \frac{K_3}{\mu_3}e^{K_3 d} & -\frac{K_3}{\mu_3}e^{-K_3 d} & \frac{K_2}{\mu_2}e^{-K_3 d} \end{pmatrix}\ . \tag{2.257}$$

A nontrivial solutions exists provided that

$$\begin{aligned} 0 &= \det M_{\mathrm{s}} \\ &= e^{-2K_3 d}\frac{\left(\frac{\mu_3}{K_3}-\frac{\mu_1}{K_1}\right)\left(\frac{\mu_3}{K_3}-\frac{\mu_2}{K_2}\right)}{\frac{\mu_1}{K_1}\frac{\mu_2}{K_2}\left(\frac{\mu_3}{K_3}\right)^2}A_{\mathrm{s}}\ , \end{aligned} \tag{2.258}$$

where

$$A_{\mathrm{s}}=e^{2K_3 d}\frac{\frac{\mu_3}{K_3}+\frac{\mu_1}{K_1}}{\frac{\mu_3}{K_3}-\frac{\mu_1}{K_1}}\frac{\frac{\mu_3}{K_3}+\frac{\mu_2}{K_2}}{\frac{\mu_3}{K_3}-\frac{\mu_2}{K_2}}-1\ . \tag{2.259}$$

Thus the angular frequencies ω associated with s-polarization surface modes can be found by solving the equation $A_{\mathrm{s}}=0$. Similarly, for the case of p-polarization, the solution for the magnetic field $\mathbf{H}$ for each layer is assumed to have the form

$$\mathbf{H}=\hat{\mathbf{y}}h_y(z)\,e^{i\boldsymbol{\kappa}\cdot\boldsymbol{\rho}}\ , \tag{2.260}$$

where [see Eq. (2.151)]

$$\frac{h_y''}{h_y}=K^2\ . \tag{2.261}$$

The electric field $\mathbf{E}$ is given by [see Eq. (2.138)]

$$\begin{aligned} \mathbf{E} &= \frac{i}{k_0\epsilon}\boldsymbol{\nabla}\times\mathbf{H} \\ &= -\frac{\left(-h_y',0,0\right)+i\left(0,0,k_x h_y\right)}{ik_0\epsilon}e^{i\boldsymbol{\kappa}\cdot\boldsymbol{\rho}}\ . \end{aligned} \tag{2.262}$$

Thus, for p-polarization the boundary condition (2.77) implies that h_y is continuous, and the boundary condition (2.78) implies that $\epsilon^{-1}h_y'$ is continuous. In a similar fashion one finds that the angular frequencies ω associated with p-polarization surface modes can be calculated by solving the equation $A_{\mathrm{p}} = 0$, where

$$A_{\mathrm{p}} = e^{2K_3 d} \frac{\frac{\epsilon_3}{K_3} + \frac{\epsilon_1}{K_1}}{\frac{\epsilon_3}{K_3} - \frac{\epsilon_1}{K_1}} \frac{\frac{\epsilon_3}{K_3} + \frac{\epsilon_2}{K_2}}{\frac{\epsilon_3}{K_3} - \frac{\epsilon_2}{K_2}} - 1 \; . \tag{2.263}$$

8. For the above discussed trilayer, the Casimir force between layers 1 and 2 is calculated below by summing over all angular frequencies ω corresponding to surface modes with either s-polarization or p-polarization [see Eqs. (2.255), (2.259) and (2.263)].
   a) The Casimir pressure (force per unit area) $P(d)$ is given by

$$P(d) = -\frac{\partial u(d)}{\partial d} \; , \tag{2.264}$$

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

$$\sum_{k_x, k_y} \rightarrow \frac{A}{4\pi^2} \int_{-\infty}^{\infty} \mathrm{d}k_x \int_{-\infty}^{\infty} \mathrm{d}k_y \rightarrow \frac{A}{2\pi} \int_0^{\infty} \mathrm{d}\kappa \; \kappa \; . \tag{2.265}$$

For a given κ the angular frequencies ω can be found by solving $A_{\mathrm{s}}(\omega, \kappa) = 0$ and $A_{\mathrm{p}}(\omega, \kappa) = 0$. This can be performed with the help of the argument theorem

$$P(d) = -\frac{\hbar}{4\pi} \frac{\partial}{\partial d} \int_0^{\infty} \mathrm{d}\kappa \; \kappa \varUpsilon(\kappa) \; , \tag{2.266}$$

where

$$\varUpsilon(\kappa) = \frac{1}{2\pi i} \oint_C \mathrm{d}\omega \; \omega \left(\frac{\partial \log A_{\mathrm{s}}}{\partial \omega} + \frac{\partial \log A_{\mathrm{p}}}{\partial \omega} \right) \; , \tag{2.267}$$

and where the integration contour C in the ω 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 $R \rightarrow \infty$ the integral along the semi-circle vanishes. By integrating along the imaginary axis and by performing integration by parts Eq. (2.266) becomes

$$P(d) = \frac{\hbar}{8\pi^2} \frac{\partial}{\partial d} \int_0^{\infty} \mathrm{d}\kappa \; \kappa \int_{-\infty}^{\infty} \mathrm{d}\varOmega \; (\log A_{\mathrm{s}} + \log A_{\mathrm{p}}) \; , \tag{2.268}$$

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

$$\begin{aligned}
P(d) &= -\frac{\hbar}{2\pi^2}\int_0^\infty \mathrm{d}\kappa\, \kappa \int_0^\infty \mathrm{d}\Omega \\
&\quad \times \left(\frac{K_3 (A_\mathrm{s}+1)}{A_\mathrm{s}} + \frac{K_3 (A_\mathrm{p}+1)}{A_\mathrm{p}} \right) \\
&= -\frac{\hbar}{\pi^2}\int_0^\infty \mathrm{d}\kappa\, \kappa \int_0^\infty \mathrm{d}\Omega\, K_3 \\
&\quad - \frac{\hbar}{2\pi^2}\int_0^\infty \mathrm{d}\kappa\, \kappa \int_0^\infty \mathrm{d}\Omega\, K_3 \left(\frac{1}{A_\mathrm{s}} + \frac{1}{A_\mathrm{p}} \right) .
\end{aligned} \tag{2.269}$$

The first term is independent on ϵ_1 and ϵ_2. After disregarding it Eq. (2.269) becomes

$$P(d) = -\frac{\hbar}{2\pi^2}\int_0^\infty \mathrm{d}\kappa\, \kappa \int_0^\infty \mathrm{d}\Omega\, K_3 \left(\frac{1}{A_\mathrm{s}} + \frac{1}{A_\mathrm{p}} \right) . \tag{2.270}$$

The variable p is defined by

$$\kappa^2 = \frac{n_3^2 \Omega^2}{c^2}\left(p^2 - 1\right) , \tag{2.271}$$

the variables $s_{1,2}$ by

$$s_{1,2} = \sqrt{p^2 - 1 + \frac{n_{1,2}^2}{n_3^2}} , \tag{2.272}$$

and the variable x by

$$x = 2K_3 d . \tag{2.273}$$

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

$$K_1 = \frac{n_3 \Omega}{c}\sqrt{p^2 - 1 + \frac{n_1^2}{n_3^2}} , \tag{2.274}$$

$$K_2 = \frac{n_3 \Omega}{c}\sqrt{p^2 - 1 + \frac{n_2^2}{n_3^2}} , \tag{2.275}$$

$$K_3 = \frac{n_3 \Omega}{c} p , \tag{2.276}$$

or

$$K_1 = \frac{x}{2pd}\sqrt{p^2 - 1 + \frac{n_1^2}{n_3^2}} \;, \tag{2.277}$$

$$K_2 = \frac{x}{2pd}\sqrt{p^2 - 1 + \frac{n_2^2}{n_3^2}} \;, \tag{2.278}$$

$$K_3 = \frac{x}{2d} \;. \tag{2.279}$$

By employing the above definitions and relations one finds that Eq. (2.269) can be expressed as [see Eqs. (2.259) and (2.263)]

$$\begin{aligned} P(d) = &-\frac{\hbar c}{32\pi^2 d^4}\int_1^\infty \frac{\mathrm{d}p}{p^2}\int_0^\infty \mathrm{d}x\; x^3 n_3^{-1} \\ &\times\left(\frac{1}{\zeta_{\mathrm{s},1}\zeta_{\mathrm{s},2}e^x - 1} + \frac{1}{\zeta_{\mathrm{p},1}\zeta_{\mathrm{p},2}e^x - 1}\right) \;, \end{aligned} \tag{2.280}$$

where

$$\zeta_{\mathrm{s},n} = \frac{\mu_n K_3 + \mu_3 K_n}{\mu_n K_3 - \mu_3 K_n} \;, \tag{2.281}$$

$$\zeta_{\mathrm{p},n} = \frac{\epsilon_n K_3 + \epsilon_3 K_n}{\epsilon_n K_3 - \epsilon_3 K_n} \;, \tag{2.282}$$

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

b) For this case [recall that $\omega = i\Omega$ and see Eqs. (2.277), (2.278) and (2.279)]

$$\zeta_{\mathrm{s},n} = \frac{1 + \sqrt{1 + x^{-2}\left(\frac{d}{d_{\mathrm{p}}}\right)^2}}{1 - \sqrt{1 + x^{-2}\left(\frac{d}{d_{\mathrm{p}}}\right)^2}} \;, \tag{2.283}$$

$$\zeta_{\mathrm{p},n} = \frac{1 + x^{-2}\left(p\frac{d}{d_{\mathrm{p}}}\right)^2 + \sqrt{1 + x^{-2}\left(\frac{d}{d_{\mathrm{p}}}\right)^2}}{1 + x^{-2}\left(p\frac{d}{d_{\mathrm{p}}}\right)^2 - \sqrt{1 + x^{-2}\left(\frac{d}{d_{\mathrm{p}}}\right)^2}} \;, \tag{2.284}$$

where

$$d_{\mathrm{p}} = \frac{c}{2\omega_{\mathrm{p}}} \;. \tag{2.285}$$

For the present case the following holds [see Eqs. (2.283) and (2.284)]

$$\frac{1}{\zeta_{\mathrm{s},n}^2 e^x - 1} + \frac{1}{\zeta_{\mathrm{p},n}^2 e^x - 1} = \frac{2}{e^x - 1}$$
$$- \frac{4xe^x}{(e^x - 1)^2}\left(1 + \frac{1}{p^2}\right)\frac{d_{\mathrm{p}}}{d} + O\left(\left(\frac{d_{\mathrm{p}}}{d}\right)^2\right) , \tag{2.286}$$

thus when $d_{\mathrm{p}} \ll d$ the Casimir force (2.280) is approximately given by

$$P(d) = P_{\mathrm{pc}}(d) + P_{\mathrm{fcc}}(d) , \tag{2.287}$$

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

$$\begin{aligned} P_{\mathrm{pc}}(d) &= -\frac{\hbar c}{16\pi^2 d^4}\int_1^\infty \frac{\mathrm{d}p}{p^2}\int_0^\infty \frac{x^3 \mathrm{d}x}{e^x - 1} \\ &= -\frac{\pi^2 \hbar c}{240 d^4} , \end{aligned} \tag{2.288}$$

represents the force in the limit of infinite conductivity, and where the term $P_{\mathrm{fcc}}(d)$, which is given by

$$P_{\mathrm{fcc}}(d) = -\frac{32}{3}\frac{d_{\mathrm{p}}}{d}P_{\mathrm{pc}}(d) , \tag{2.289}$$

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

$$\int_0^\infty \frac{x^3 \mathrm{d}x}{e^x - 1} = \frac{\pi^4}{15} , \tag{2.290}$$

$$\int_1^\infty \frac{\mathrm{d}p}{p^2} = 1 , \tag{2.291}$$

$$\int_0^\infty \frac{x^4 e^x \mathrm{d}x}{(e^x - 1)^2} = \frac{4\pi^4}{15} , \tag{2.292}$$

$$\int_1^\infty \frac{\left(1 + \frac{1}{p^2}\right)\mathrm{d}p}{p^2} = \frac{4}{3} . \tag{2.293}$$

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

$$\begin{aligned} \mathbf{E}' &= \mathbf{E}_\parallel + \gamma\left(\mathbf{E}_\perp + \boldsymbol{\beta}\times\mathbf{B}_\perp\right) \\ &= \gamma\mathbf{E} + (1-\gamma)\left(\mathbf{E}\cdot\hat{\boldsymbol{\beta}}\right)\hat{\boldsymbol{\beta}} + \gamma\beta\hat{\boldsymbol{\beta}}\times\mathbf{B} , \end{aligned} \tag{2.294}$$

where $\gamma = 1/\sqrt{1-\beta^2}$ and $\mathbf{B} = \hat{\mathbf{z}} \times \mathbf{E} = (-E_y, E_x, 0)$ [see Eq. (2.139)], and thus for the case where $\hat{\boldsymbol{\beta}} = (\sin\theta, 0, \cos\theta)$ one has

$$E'_x = \left(1 + (\gamma - 1)\cos^2\theta - \beta\gamma\cos\theta\right) E_x \; , \tag{2.295}$$

$$E'_y = \gamma\left(1 - \beta\cos\theta\right) E_y \; , \tag{2.296}$$

$$E'_z = \left((1-\gamma)\cos\theta + \beta\gamma\right) E_x \sin\theta \; . \tag{2.297}$$

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

$$\hat{\mathbf{n}}' = \frac{\hat{\mathbf{n}} - \gamma\left(1 - \frac{\gamma}{1+\gamma}\left(\boldsymbol{\beta}\cdot\hat{\mathbf{n}}\right)\right)\boldsymbol{\beta}}{\gamma\left(1 - \left(\boldsymbol{\beta}\cdot\hat{\mathbf{n}}\right)\right)} \; , \tag{2.298}$$

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

$$\tan\alpha = -\frac{\left((1-\gamma)\cos\theta + \beta\gamma\right)\sin\theta}{1 + (\gamma-1)\cos^2\theta - \beta\gamma\cos\theta} \; , \tag{2.299}$$

and the following holds $\tan\alpha = -E'_z/E'_x$ [see Eqs. (2.295) and (2.297)]. Rotation of $\mathbf{E}'$ about the y axis by an angle α leads to (note that the rotation preserves both the y component and the length)

$$\begin{pmatrix} E'_x \\ E'_y \\ E'_z \end{pmatrix} \rightarrow \begin{pmatrix} E'_{x\mathrm{R}} \\ E'_{y\mathrm{R}} \\ E'_{z\mathrm{R}} \end{pmatrix} = \begin{pmatrix} \sqrt{E'^2_x + E'^2_z} \\ E'_y \\ 0 \end{pmatrix} \; , \tag{2.300}$$

and thus

$$\begin{pmatrix} E'_{x\mathrm{R}} \\ E'_{y\mathrm{R}} \end{pmatrix} = \gamma\left(1 - \beta\cos\theta\right)\begin{pmatrix} E_x \\ E_y \end{pmatrix} \; , \tag{2.301}$$

and therefore the Stokes parameters are transformed according to

$$\begin{pmatrix} S'_0 \\ S'_1 \\ S'_2 \\ S'_3 \end{pmatrix} = \gamma^2\left(1 - \beta\cos\theta\right)^2\begin{pmatrix} S_0 \\ S_1 \\ S_2 \\ S_3 \end{pmatrix} \; . \tag{2.302}$$

As can be seen from Eq. (1.183), the factor $\gamma\left(1 - \beta\cos\theta\right)$ is the frequency ratio ω'/ω due to the Doppler effect.

10. In the presence of an applied electric filed given by $\mathbf{E} = (E_x, E_y, 0)\, e^{-i\omega t}$ [see Eq. (2.127)], where E_x, E_y and ω are constants, the mechanical equation of motion is given by

$$m\left(\ddot{\mathbf{r}} + \gamma\dot{\mathbf{r}} + \omega_0^2\mathbf{r}\right) = -e\mathbf{E} - \frac{e}{c}\dot{\mathbf{r}} \times \mathbf{H} \; . \tag{2.303}$$

a) The vector $\mathbf{r}$ can be expressed as $\mathbf{r} = (r_x, r_y, 0)\, e^{-i\omega t} = (r_- \hat{\mathbf{e}}_+ + r_+ \hat{\mathbf{e}}_-)\, e^{-i\omega t}$, where $r_\pm = (r_x \pm i r_y)/\sqrt{2}$ and $\hat{\mathbf{e}}_\pm = (\hat{\mathbf{x}} \pm i \hat{\mathbf{y}})/\sqrt{2}$. With the help of the identity $\hat{\mathbf{e}}_\pm \times \hat{\mathbf{z}} = \pm i \hat{\mathbf{e}}_\pm$ Eq. (2.303) yields in steady state

$$\begin{aligned}&\left(\omega_0^2 - \omega^2 - i\omega\gamma\right)(r_- \hat{\mathbf{e}}_+ + r_+ \hat{\mathbf{e}}_-)\\&+ \omega\omega_{\mathrm{H}} (r_- \hat{\mathbf{e}}_+ - r_+ \hat{\mathbf{e}}_-)\\&= -\frac{e}{m} (E_- \hat{\mathbf{e}}_+ + E_+ \hat{\mathbf{e}}_-)\ ,\end{aligned} \tag{2.304}$$

where $(E_x, E_y, 0) = E_- \hat{\mathbf{e}}_+ + E_+ \hat{\mathbf{e}}_-$ and the cyclotron frequency ω_{H} is given by

$$\omega_{\mathrm{H}} = \frac{e H_0}{mc}\ , \tag{2.305}$$

and thus

$$\frac{r_\pm}{E_\pm} = -\frac{\frac{e}{m}}{\omega_0^2 - \omega^2 - i\omega\gamma \mp \omega\omega_{\mathrm{H}}}\ , \tag{2.306}$$

or (recall that it is assumed that $\gamma \ll |\omega - \omega_0|$ and $\omega_{\mathrm{H}} \ll \omega_0$)

$$\frac{r_\pm}{E_\pm} = -\frac{\frac{e}{m}\left(1 \pm \frac{\omega\omega_{\mathrm{H}}}{\omega_0^2 - \omega^2}\right)}{\omega_0^2 - \omega^2} + O\left(\omega_{\mathrm{H}}^2\right)\ . \tag{2.307}$$

With the help of the relations $\mathbf{D} = \mathbf{E} + 4\pi\mathbf{P}$ and $\mathbf{P} = \chi_{\mathrm{e}}\mathbf{E}$ [see Eqs. (2.17) and (2.86)] one finds that the permittivity $\epsilon = 1 + 4\pi\chi_{\mathrm{e}}$ corresponding to circular polarization $\hat{\mathbf{e}}_\pm$ is given by

$$\epsilon_\pm = 1 + \frac{\omega_{\mathrm{p}}^2\left(1 \pm \frac{\omega\omega_{\mathrm{H}}}{\omega_0^2 - \omega^2}\right)}{\omega_0^2 - \omega^2} + O\left(\omega_{\mathrm{H}}^2\right)\ , \tag{2.308}$$

where

$$\omega_{\mathrm{p}} = \sqrt{\frac{4\pi N e^2}{m}}\ , \tag{2.309}$$

is the plasma frequency, or

$$\epsilon_\pm = n_0^2\left(1 \pm \frac{\omega_{\mathrm{p}}^2 \omega\omega_{\mathrm{H}}}{n_0^2\left(\omega_0^2 - \omega^2\right)^2}\right) + O\left(\omega_{\mathrm{H}}^2\right)\ , \tag{2.310}$$

where

$$n_0 = \left(1 + \frac{\omega_{\mathrm{p}}^2}{\omega_0^2 - \omega^2}\right)^{1/2}\ , \tag{2.311}$$

and the corresponding indices of refraction are given by $n_\pm = \epsilon_\pm^{1/2}$.

b) According to its definition, the Verdet constant is given by

$$V = \frac{\omega (n_+ - n_-)}{cH_0} . \tag{2.312}$$

hence

$$V = \frac{e}{mc^2 n_0} \frac{\omega_{\mathrm{p}}^2 \omega^2}{\left(\omega_0^2 - \omega^2\right)^2} . \tag{2.313}$$

c) The magnetization $\mathbf{M}$ is given by [see Eq. (2.137) and note that $\hat{\mathbf{e}}_\pm^* = \hat{\mathbf{e}}_\mp$ and $\hat{\mathbf{e}}_- \times \hat{\mathbf{e}}_+ = i\hat{\mathbf{z}}$]

$$\begin{aligned} \mathbf{M} &= \frac{\pi e N}{2c} \mathbf{r} \times \dot{\mathbf{r}}^* \\ &= \frac{i\pi e N \omega}{2c} (r_- \hat{\mathbf{e}}_+ + r_+ \hat{\mathbf{e}}_-) \times \left(r_-^* \hat{\mathbf{e}}_- + r_+^* \hat{\mathbf{e}}_+\right) \\ &= \frac{\pi e N \omega}{2c} \left(|r_-|^2 - |r_+|^2\right) \hat{\mathbf{z}} , \end{aligned} \tag{2.314}$$

(2.315)

thus in the limit of $\gamma \to 0$ and $\omega_{\mathrm{H}} \to 0$ [see Eq. (2.306)]

$$\begin{aligned} \mathbf{M} &= \frac{\pi e N \omega}{2c} \left(\frac{\frac{e}{m}}{\omega_0^2 - \omega^2}\right)^2 \left(|E_-|^2 - |E_+|^2\right) \hat{\mathbf{z}} \\ &= \frac{c n_0 V}{8\omega} \left(|E_-|^2 - |E_+|^2\right) \hat{\mathbf{z}} . \end{aligned} \tag{2.316}$$

11. The following holds [see Eqs. (1.21), (2.32) and (2.49)]

$$\eta \hat{F} \eta \hat{F} = \Lambda^{-1} \eta \hat{F}' \eta \hat{F}' \Lambda, \tag{2.317}$$

$$\eta \hat{G} \eta \hat{G} = \Lambda^{-1} \eta \hat{G}' \eta \hat{G}' \Lambda, \tag{2.318}$$

and [see Eqs. (1.14), (2.29) and (2.46)]

$$\frac{1}{2} \operatorname{Tr} \left(\eta \hat{F} \eta \hat{F}\right) = \mathbf{E}^2 - \mathbf{B}^2, \tag{2.319}$$

$$\frac{1}{2} \operatorname{Tr} \left(\eta \hat{G} \eta \hat{G}\right) = \mathbf{D}^2 - \mathbf{H}^2. \tag{2.320}$$

In general, for any two square matrices M_1 and M_2 the following holds $\operatorname{Tr}(M_1 M_2) = \operatorname{Tr}(M_2 M_1)$, and thus

$$\operatorname{Tr} \left(\eta \hat{F} \eta \hat{F}\right) = \operatorname{Tr} \left(\eta \hat{F}' \eta \hat{F}'\right) , \tag{2.321}$$

$$\operatorname{Tr} \left(\eta \hat{G} \eta \hat{G}\right) = \operatorname{Tr} \left(\eta \hat{G}' \eta \hat{G}'\right) . \tag{2.322}$$

i.e. both $\mathbf{E}^2 - \mathbf{B}^2$ and $\mathbf{D}^2 - \mathbf{H}^2$ are Lorentz invariant. The following holds [see Eqs. (2.35), (2.36) and (3.315)]

$$
\begin{aligned}
\mathbf{E}\cdot\mathbf{B} &= \mathbf{E}'_{\|}\cdot\mathbf{B}'_{\|}+\gamma^2\left(\mathbf{E}'_{\perp}-\boldsymbol{\beta}\times\mathbf{B}'_{\perp}\right)\cdot\left(\mathbf{B}'_{\perp}+\boldsymbol{\beta}\times\mathbf{E}'_{\perp}\right) \\
&= \mathbf{E}'_{\|}\cdot\mathbf{B}'_{\|}+\gamma^2\left(1-\beta^2\right)\mathbf{E}'_{\perp}\cdot\mathbf{B}'_{\perp}+\gamma^2\left(\mathbf{E}'_{\perp}\cdot\left(\boldsymbol{\beta}\times\mathbf{E}'_{\perp}\right)-\left(\boldsymbol{\beta}\times\mathbf{B}'_{\perp}\right)\cdot\mathbf{B}'_{\perp}\right) \\
&= \mathbf{E}'\cdot\mathbf{B}' \;,
\end{aligned}
\tag{2.323}
$$

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. (2.104) and (2.105)]. Moreover, with the help of the general identity

$$
\left(\mathbf{A}\times\mathbf{B}\right)\cdot\mathbf{C}=\left(\mathbf{B}\times\mathbf{C}\right)\cdot\mathbf{A}=\left(\mathbf{C}\times\mathbf{A}\right)\cdot\mathbf{B}\;, \tag{2.324}
$$

one finds that [see Eqs. (2.35), (2.36), (2.104) and (2.105)]

$$
\begin{aligned}
\mathbf{E}\cdot\mathbf{D}-\mathbf{B}\cdot\mathbf{H} &= \mathbf{E}'_{\|}\cdot\mathbf{D}'_{\|}-\mathbf{B}'_{\|}\cdot\mathbf{H}'_{\|} \\
&\quad+\gamma^2\left(\mathbf{E}'_{\perp}-\boldsymbol{\beta}\times\mathbf{B}'_{\perp}\right)\cdot\left(\mathbf{D}'_{\perp}-\boldsymbol{\beta}\times\mathbf{H}'_{\perp}\right) \\
&\quad-\gamma^2\left(\mathbf{B}'_{\perp}+\boldsymbol{\beta}\times\mathbf{E}'_{\perp}\right)\cdot\left(\mathbf{H}'_{\perp}+\boldsymbol{\beta}\times\mathbf{D}'_{\perp}\right) \\
&= \mathbf{E}'_{\|}\cdot\mathbf{D}'_{\|}-\mathbf{B}'_{\|}\cdot\mathbf{H}'_{\|}+\gamma^2\left(1-\beta^2\right)\left(\mathbf{E}'_{\perp}\cdot\mathbf{D}'_{\perp}-\mathbf{B}'_{\perp}\cdot\mathbf{H}'_{\perp}\right) \\
&= \mathbf{E}'\cdot\mathbf{D}'-\mathbf{B}'\cdot\mathbf{H}'\;,
\end{aligned}
\tag{2.325}
$$

and thus $\mathbf{E}\cdot\mathbf{D}-\mathbf{B}\cdot\mathbf{H}$ is Lorentz invariant.

12. Both $\mathbf{E}\cdot\mathbf{B}$ and $\mathbf{E}\cdot\mathbf{E}-\mathbf{B}\cdot\mathbf{B}$ are Lorentz invariant, and thus the electric field $\mathbf{E}$ in the inertial frame S vanishes for any position $\mathbf{x}$ and at any time t.
13. In the Lorenz gauge the Maxwell's equations in vacuum can be expressed as [see Eqs. (2.62) and (2.63)]

$$
\Box^2 A=-\frac{4\pi}{c}J\;, \tag{2.326}
$$

where $\Box^2=-c^{-2}\partial^2/\partial t^2+\boldsymbol{\nabla}^2$ is the D'Alembertian operator, $A=\left(\phi,A_1,A_2,A_3\right)^{\mathrm{T}}$ is the potential 4-vector and $J=\left(c\rho,J_1,J_2,J_3\right)^{\mathrm{T}}$ is the current 4-vector. Both $A\left(t,\mathbf{x}\right)$ and $J\left(t,\mathbf{x}\right)$ can be Fourier expanded (in time only) as

$$
A\left(t,\mathbf{x}\right)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}\mathrm{d}\omega\;A\left(\omega,\mathbf{x}\right)e^{i\omega t}\;, \tag{2.327}
$$

$$
J\left(t,\mathbf{x}\right)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}\mathrm{d}\omega\;J\left(\omega,\mathbf{x}\right)e^{i\omega t}\;. \tag{2.328}
$$

Substituting into Eq. (2.326) yields

$$
\left(\frac{\omega^2}{c^2}+\boldsymbol{\nabla}^2\right)A\left(\omega,\mathbf{x}\right)=-\frac{4\pi}{c}J\left(\omega,\mathbf{x}\right)\;. \tag{2.329}
$$

As is shown below in chapter 5 [see Eqs. (5.84) and (5.95)], the following holds

$$\left(k^2 + \boldsymbol{\nabla}^2\right) g\left(\mathbf{x} - \mathbf{x}'\right) = \delta\left(\mathbf{x} - \mathbf{x}'\right) \ , \tag{2.330}$$

where k is a constant, and the so-called Green function $g\left(\mathbf{x} - \mathbf{x}'\right)$ is given by

$$g\left(\mathbf{x} - \mathbf{x}'\right) = -\frac{e^{\pm ik\left|\mathbf{x}-\mathbf{x}'\right|}}{4\pi\left|\mathbf{x} - \mathbf{x}'\right|} \ . \tag{2.331}$$

Thus the solution of Eq. (2.329) can be expressed as

$$A\left(\omega, \mathbf{x}\right) = \int \mathrm{d}^3\mathbf{x}' \ \frac{e^{\pm i\frac{\omega}{c}\left|\mathbf{x}-\mathbf{x}'\right|}}{c\left|\mathbf{x} - \mathbf{x}'\right|} J\left(\omega, \mathbf{x}'\right) \ . \tag{2.332}$$

Applying the inverse Fourier transform in time leads to [see Eq. (5.6)]

$$\begin{aligned} A\left(t, \mathbf{x}\right) &= \int \mathrm{d}^3\mathbf{x}' \int_{-\infty}^{\infty} \mathrm{d}t' \ \frac{J\left(t', \mathbf{x}'\right)}{c\left|\mathbf{x} - \mathbf{x}'\right|} \underbrace{\frac{1}{2\pi}\int_{-\infty}^{\infty} \mathrm{d}\omega \ e^{i\omega\left(t \pm \frac{\left|\mathbf{x}-\mathbf{x}'\right|}{c} - t'\right)}}_{=\delta\left(t \pm \frac{\left|\mathbf{x}-\mathbf{x}'\right|}{c} - t'\right)} \\ &= \int \mathrm{d}^3\mathbf{x}' \ \frac{J\left(t \pm \frac{\left|\mathbf{x}-\mathbf{x}'\right|}{c}, \mathbf{x}'\right)}{c\left|\mathbf{x} - \mathbf{x}'\right|} \ . \end{aligned} \tag{2.333}$$

Due to the principle of causality, the solution with the + sign is rejected. The potential with the minus sign, which is given by Eq. (2.185), 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 $t - \left|\mathbf{x} - \mathbf{x}'\right|/c$ and location $\mathbf{x}'$ affecting the value of A at time t and location $\mathbf{x}$.

14. For the current case Eq. (2.185) yields [see Eq. (1.112)]

$$A\left(t, \mathbf{x}\right) = \frac{\left(\frac{q}{r_0}, -\frac{q\omega}{c}\sin\omega t_{\mathrm{r}}, \frac{q\omega}{c}\cos\omega t_{\mathrm{r}}, 0\right)^{\mathrm{T}}}{\sqrt{1 + \left(\frac{z}{r_0}\right)^2}} \ , \tag{2.334}$$

where the retarded time t_{r} is given by

$$t_{\mathrm{r}} = t - \frac{\sqrt{r_0^2 + z^2}}{c} \ . \tag{2.335}$$

15. For a general point particle of charge q the current 4-vector is given by [see Eqs. (1.37), (1.110), (1.111) and (1.112)]

$$J = q\left(c, \dot{\mathbf{x}}_{\mathrm{p}}\right)^{\mathrm{T}} \delta\left(\mathbf{x} - \mathbf{x}_{\mathrm{p}}\right) \ , \tag{2.336}$$

where $\mathbf{x}_{\mathrm{p}}$ and $\dot{\mathbf{x}}_{\mathrm{p}}$ are the position and velocity, respectively, of the point particle, hence the Leinard–Wiechert potential at time t and position $\mathbf{x} = (x_1, x_2, x_3)$ is given by [see Eq. (2.185)]

$$A(t,\mathbf{x}) = \int \mathrm{d}^3\mathbf{x}' \; \frac{q\left(1, \boldsymbol{\beta}\left(t'_{\mathrm{r}}\right)\right)^{\mathrm{T}} \delta\left(\mathbf{x}' - \mathbf{x}_{\mathrm{p}}\left(t'_{\mathrm{r}}\right)\right)}{\left|\mathbf{x} - \mathbf{x}'\right|} \;, \tag{2.337}$$

where the retarded time t'_{r} is given by $t'_{\mathrm{r}} = t - t_{\mathrm{d}}$, the delay time t_{d} is given by $t_{\mathrm{d}} = \left|\mathbf{x} - \mathbf{x}'\right|/c$, $\gamma = 1/\sqrt{1-\beta^2}$, $\beta = \left|\dot{\mathbf{x}}_{\mathrm{p}}\right|/c$, and $\boldsymbol{\beta} = \dot{\mathbf{x}}_{\mathrm{p}}/c$. A delta function in time and time integration can be added

$$A(t,\mathbf{x}) = \int \mathrm{d}^3\mathbf{x}' \int \mathrm{d}t' \; \frac{q\left(1, \boldsymbol{\beta}\left(t'\right)\right)^{\mathrm{T}} \delta\left(\mathbf{x}' - \mathbf{x}_{\mathrm{p}}\left(t'\right)\right)\delta\left(t' - t'_{\mathrm{r}}\right)}{\left|\mathbf{x} - \mathbf{x}'\right|} \;. \tag{2.338}$$

Integration over space yields

$$A(t,\mathbf{x}) = \int \mathrm{d}t' \; \frac{q\left(1, \boldsymbol{\beta}\left(t'\right)\right)^{\mathrm{T}} \delta\left(f\left(t'\right)\right)}{\left|\mathbf{x} - \mathbf{x}_{\mathrm{p}}\left(t'\right)\right|} \;, \tag{2.339}$$

where the function $f(t')$ is given by

$$\begin{aligned} f\left(t'\right) &= t' - t'_{\mathrm{r}} \\ &= t' - t + \frac{\left|\mathbf{x} - \mathbf{x}_{\mathrm{p}}\left(t'\right)\right|}{c} \;. \end{aligned} \tag{2.340}$$

For the case where the function $f(t')$ has a single zero at t_{r}, the following holds

$$\begin{aligned} \delta\left(f\left(t'\right)\right) &= \frac{\delta\left(t' - t_{\mathrm{r}}\right)}{\left|\frac{\mathrm{d}f(t')}{\mathrm{d}t'}\right|} \\ &= \frac{\delta\left(t' - t_{\mathrm{r}}\right)}{1 - \hat{\mathbf{n}}_{\mathrm{p}} \cdot \boldsymbol{\beta}} \;, \end{aligned} \tag{2.341}$$

where the unit vector $\hat{\mathbf{n}}_{\mathrm{p}}$ is given by

$$\hat{\mathbf{n}}_{\mathrm{p}}\left(t'\right) = \frac{\mathbf{x} - \mathbf{x}_{\mathrm{p}}\left(t'\right)}{\left|\mathbf{x} - \mathbf{x}_{\mathrm{p}}\left(t'\right)\right|} \;, \tag{2.342}$$

and thus $A(t,\mathbf{x})$ can be expressed as

$$A(t,\mathbf{x}) = \frac{q\left(1, \boldsymbol{\beta}\left(t_{\mathrm{r}}\right)\right)^{\mathrm{T}}}{\left(1 - \hat{\mathbf{n}}_{\mathrm{p}}\left(t_{\mathrm{r}}\right) \cdot \boldsymbol{\beta}\left(t_{\mathrm{r}}\right)\right)\left|\mathbf{x} - \mathbf{x}_{\mathrm{p}}\left(t_{\mathrm{r}}\right)\right|} \;. \tag{2.343}$$

For the current case, the particle position is given by $\mathbf{x}_{\mathrm{p}}(t) = (ut, 0, 0)$, and thus the condition $f(t') = 0$ yields [see Eq. (2.340)]

$$0 = t_{\mathrm{r}} - t + \frac{|\mathbf{x} - (ut_{\mathrm{r}}, 0, 0)|}{c} \ , \tag{2.344}$$

hence

$$\begin{aligned} t_{\mathrm{r}} - t &= -\frac{\gamma^2 \left(\beta (x_1 - ut) \pm \sqrt{(x_1 - ut)^2 + \frac{x_2^2 + x_3^2}{\gamma^2}} \right)}{c} \\ &= -\frac{\gamma^2 \left(\frac{\mathbf{x}_{\mathrm{r}} \cdot \dot{\mathbf{x}}_{\mathrm{p}}}{c} \pm \sqrt{\left(\frac{\mathbf{x}_{\mathrm{r}} \cdot \dot{\mathbf{x}}_{\mathrm{p}}}{c} \right)^2 + \frac{|\mathbf{x}_{\mathrm{r}}|^2}{\gamma^2}} \right)}{c} \ , \end{aligned} \tag{2.345}$$

where

$$\mathbf{x}_{\mathrm{r}} = \mathbf{x} - \mathbf{x}_{\mathrm{p}} (t) \ . \tag{2.346}$$

For the solution with the plus sign, which satisfies the causality condition, Eq. (2.343) yields [see Eq. (2.340), recall that $f(t_{\mathrm{r}}) = 0$, and note that $\mathbf{x} - \mathbf{x}_{\mathrm{p}} (t_{\mathrm{r}}) = \mathbf{x}_{\mathrm{r}} - (u (t_{\mathrm{r}} - t), 0, 0)$]

$$\begin{aligned} A(t, \mathbf{x}) &= \frac{q (1, \boldsymbol{\beta})^{\mathrm{T}}}{|\mathbf{x} - \mathbf{x}_{\mathrm{p}} (t_{\mathrm{r}})| - (\mathbf{x} - \mathbf{x}_{\mathrm{p}} (t')) \cdot \boldsymbol{\beta}} \\ &= \frac{q (1, \boldsymbol{\beta})^{\mathrm{T}}}{-c (t_{\mathrm{r}} - t) - (\mathbf{x}_{\mathrm{r}} - (u (t_{\mathrm{r}} - t), 0, 0)) \cdot \boldsymbol{\beta}} \\ &= \frac{q (1, \boldsymbol{\beta})^{\mathrm{T}}}{-\frac{c(t_{\mathrm{r}} - t)}{\gamma^2} - \mathbf{x}_{\mathrm{r}} \cdot \boldsymbol{\beta}} \\ &= \frac{\gamma q (1, \boldsymbol{\beta})^{\mathrm{T}}}{r_0} \ , \end{aligned} \tag{2.347}$$

where

$$r_0 = \sqrt{|\mathbf{x}_{\mathrm{r}}|^2 + \gamma^2 (\mathbf{x}_{\mathrm{r}} \cdot \boldsymbol{\beta})^2} = \sqrt{\gamma^2 (x_1 - ut)^2 + x_2^2 + x_3^2} \ . \tag{2.348}$$

Hence the electric field is given by [see Eq. (2.25) and recall that $A = (\phi, A_1, A_2, A_3)^{\mathrm{T}}$]

$$\begin{aligned} \mathbf{E} &= -\boldsymbol{\nabla} \phi - \frac{1}{c} \frac{\partial \mathbf{A}}{\partial t} \\ &= -\gamma q \left(\boldsymbol{\nabla} \left(\frac{1}{r_0} \right) + \frac{1}{c} \frac{\partial}{\partial t} \left(\frac{1}{r_0} \right) \boldsymbol{\beta} \right) \\ &= \frac{\gamma q \mathbf{x}_{\mathrm{r}}}{r_0^3} \ , \end{aligned} \tag{2.349}$$

and the magnetic field is given by [see Eq. (2.23)]

$$\begin{aligned}\mathbf{B} &= \boldsymbol{\nabla} \times \mathbf{A} \\ &= \gamma q \boldsymbol{\nabla} \times \frac{\boldsymbol{\beta}}{r_0} \\ &= \gamma q \left(\boldsymbol{\nabla} \frac{1}{r_0}\right) \times \boldsymbol{\beta} \\ &= -\frac{\gamma q \left(\left(\gamma^2 (x_1 - ut), x_2, x_3\right)\right)}{r_0^3} \times \boldsymbol{\beta} \\ &= \gamma q \frac{\boldsymbol{\beta} \times \mathbf{x}_{\mathrm{r}}}{r_0^3} \;,\end{aligned} \tag{2.350}$$

in agreement with Eqs. (2.193) and (2.194), respectively.

16. In the dipole approximation (i.e. when $|\mathbf{x}'| \ll |\mathbf{x}|$) the Leinard–Wiechert potential 4-vector A becomes [see Eq. (2.185)]

$$\begin{aligned}A(t, \mathbf{x}) &= \int \mathrm{d}^3\mathbf{x}' \; \frac{J\left(t - \frac{|\mathbf{x}-\mathbf{x}'|}{c}, \mathbf{x}'\right)}{c\,|\mathbf{x}-\mathbf{x}'|} \\ &\simeq \frac{1}{c\,|\mathbf{x}|} \int \mathrm{d}^3\mathbf{x}' \; J\left(t - \frac{|\mathbf{x}|}{c}, \mathbf{x}'\right) \;.\end{aligned} \tag{2.351}$$

With the help of the continuity equation (1.117) one finds that [see Eqs. (1.112) and (2.186)]

$$\dot{\mathbf{p}} = -\int \mathrm{d}^3\mathbf{x}' \; \mathbf{x}' \left(\boldsymbol{\nabla} \cdot \mathbf{J}\right) \;, \tag{2.352}$$

where overdot denotes a derivative with respect to time and $\mathbf{J}$ is the 3-vector current density. Integration by parts yields

$$\dot{\mathbf{p}} = \int \mathrm{d}^3\mathbf{x}' \; \mathbf{J} \;, \tag{2.353}$$

and thus in this approximation the 3-vector potential $\mathbf{A}$ becomes

$$\mathbf{A}(t, \mathbf{x}) = \frac{\dot{\mathbf{p}}\left(t - \frac{|\mathbf{x}|}{c}\right)}{c\,|\mathbf{x}|} \;. \tag{2.354}$$

The scalar potential ϕ can be evaluated using the Lorenz gauge condition (2.61), which for the current case becomes [see Eq. (2.354)]

$$\frac{\partial \phi}{\partial t} = -\boldsymbol{\nabla} \cdot \frac{\dot{\mathbf{p}}\left(t - \frac{|\mathbf{x}|}{c}\right)}{|\mathbf{x}|} = \frac{\dot{\mathbf{p}}\left(t - \frac{|\mathbf{x}|}{c}\right) + \frac{|\mathbf{x}|}{c}\ddot{\mathbf{p}}\left(t - \frac{|\mathbf{x}|}{c}\right)}{|\mathbf{x}|^2} \cdot \hat{\mathbf{x}} \;, \tag{2.355}$$

where $\hat{\mathbf{x}} = \mathbf{x}/|\mathbf{x}|$ is a unit vector parallel to $\hat{\mathbf{x}}$, and thus by integration one obtains

$$\phi = \phi_0 + \frac{\mathbf{p}\left(t - \frac{|\mathbf{x}|}{c}\right) + \frac{|\mathbf{x}|}{c}\dot{\mathbf{p}}\left(t - \frac{|\mathbf{x}|}{c}\right)}{|\mathbf{x}|^2} \cdot \hat{\mathbf{x}} \,, \tag{2.356}$$

where the electrostatic term ϕ_0 is given by [see Eq. (2.351)]

$$\phi_0 = \frac{1}{|\mathbf{x}|}\int \mathrm{d}^3\mathbf{x}' \, \rho \,. \tag{2.357}$$

The magnetic $\mathbf{B}$ and electric $\mathbf{E}$ fields can be calculated using Eqs. (2.23) and (2.25). In the far field limit only the terms of lowest nonvanishing order in $|\mathbf{x}|^{-1}$ are kept, and thus [see Eq. (2.150)]

$$\mathbf{B} = \boldsymbol{\nabla} \times \mathbf{A} = -\frac{\hat{\mathbf{x}} \times \ddot{\mathbf{p}}_{\mathrm{r}}}{c^2 |\mathbf{x}|} \,, \tag{2.358}$$

and

$$\mathbf{E} = -\boldsymbol{\nabla}\phi - \frac{1}{c}\frac{\partial \mathbf{A}}{\partial t} = \frac{(\ddot{\mathbf{p}}_{\mathrm{r}} \cdot \hat{\mathbf{x}})\,\hat{\mathbf{x}} - \ddot{\mathbf{p}}_{\mathrm{r}}}{c^2 |\mathbf{x}|} \,, \tag{2.359}$$

or [see Eq. (1.96)]

$$\mathbf{E} = \frac{\hat{\mathbf{x}} \times (\hat{\mathbf{x}} \times \ddot{\mathbf{p}}_{\mathrm{r}})}{c^2 |\mathbf{x}|} \,, \tag{2.360}$$

where $\ddot{\mathbf{p}}_{\mathrm{r}}$ denotes the value of $\ddot{\mathbf{p}}$ at the retarded time $t - |\mathbf{x}|/c$. The Poynting vector (2.93) is given by [see Eq. (3.65)]

$$\mathbf{S} = \frac{c}{4\pi}\mathbf{E} \times \mathbf{B} = \frac{|\hat{\mathbf{x}} \times \ddot{\mathbf{p}}_{\mathrm{r}}|^2 \,\hat{\mathbf{x}}}{4\pi c^3 |\mathbf{x}|^2} \,. \tag{2.361}$$

The total radiated power P is calculated by surface integration over a sphere [see Eq. (2.92)]

$$P = \int_S \mathbf{S} \cdot \mathrm{d}\mathbf{s} = \frac{|\ddot{\mathbf{p}}_{\mathrm{r}}|^2}{4\pi c^3} \int_{-1}^{1} \mathrm{d}\cos\theta \, \sin^2\theta \int_0^{2\pi} \mathrm{d}\varphi \,, \tag{2.362}$$

thus

$$P = \frac{2\,|\ddot{\mathbf{p}}_{\mathrm{r}}|^2}{3c^3} \,. \tag{2.363}$$

3. Geometrical Optics

In the theory of geometrical optics the Maxwell's equations are simplified based on the assumption that the characteristic wavelength λ 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. (3.195)]

$$\left(\frac{n^2(\mathbf{r})}{c^2}\frac{\partial^2}{\partial t^2} - \boldsymbol{\nabla}^2\right)U(\mathbf{r},t) = 0\ . \tag{3.1}$$

By substituting a solution having the form

$$U(\mathbf{r},t) = u(\mathbf{r})\,e^{-i\omega t}\ , \tag{3.2}$$

into Eq. (3.1) one finds that $u(\mathbf{r})$ satisfies the following Helmholtz equation [see Eq. (2.151)]

$$\left(\boldsymbol{\nabla}^2 + n^2 k_0^2\right)u(\mathbf{r}) = 0\ , \tag{3.3}$$

where

$$k_0 = \frac{\omega}{c}\ . \tag{3.4}$$

Consider a solution for $u(\mathbf{r})$ having the form

$$u(\mathbf{r}) = u_{\mathrm{E}}(\mathbf{r})\,e^{ik_0\psi(\mathbf{r})}\ , \tag{3.5}$$

where $u_{\mathrm{E}}(\mathbf{r})$ is expressed as an asymptotic expansion in powers of $1/k_0$

$$u_{\mathrm{E}}(\mathbf{r}) = u_0(\mathbf{r}) + \frac{1}{k_0}u_1(\mathbf{r}) + \frac{1}{k_0^2}u_2(\mathbf{r}) + \cdots\ , \tag{3.6}$$

and where the real function $\psi(\mathbf{r})$ is called the *eikonal* (image in greek). In geometrical optics the parameter $1/k_0$ is assumed small.

Claim. To first order in $1/k_0$ the following holds

$$(\boldsymbol{\nabla}\psi)^2 = n^2 , \tag{3.7}$$

and

$$u_0\boldsymbol{\nabla}^2\psi + 2(\boldsymbol{\nabla}\psi)\cdot(\boldsymbol{\nabla}u_0) = 0 . \tag{3.8}$$

Proof. By substituting Eq. (3.5) into Eq. (3.3) and employing the general identity

$$\boldsymbol{\nabla}^2(fg) = f\boldsymbol{\nabla}^2 g + g\boldsymbol{\nabla}^2 f + 2(\boldsymbol{\nabla}f)\cdot(\boldsymbol{\nabla}g) , \tag{3.9}$$

one obtains

$$k_0^{-2}\boldsymbol{\nabla}^2 u_{\mathrm{E}} + ik_0^{-1}u_{\mathrm{E}}\left(\boldsymbol{\nabla}^2\psi + ik_0(\boldsymbol{\nabla}\psi)^2\right) + 2ik_0^{-1}(\boldsymbol{\nabla}\psi)\cdot(\boldsymbol{\nabla}u_{\mathrm{E}}) + n^2u_{\mathrm{E}} = 0 . \tag{3.10}$$

Collecting all terms of zeroth order in $1/k_0$ yields

$$\left[n^2 - (\boldsymbol{\nabla}\psi)^2\right]u_0 = 0 , \tag{3.11}$$

and collecting all terms of 1'st order in $1/k_0$ yields

$$i\left[u_0\boldsymbol{\nabla}^2\psi + 2(\boldsymbol{\nabla}\psi)\cdot(\boldsymbol{\nabla}u_0)\right] + \left[n^2 - (\boldsymbol{\nabla}\psi)^2\right]u_1 = 0 . \tag{3.12}$$

Unless $u_0(\mathbf{r})$ 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

$$u_0^2\boldsymbol{\nabla}^2\psi + (\boldsymbol{\nabla}\psi)\cdot\left(\boldsymbol{\nabla}u_0^2\right) = 0 , \tag{3.13}$$

thus Eq. (3.8) can be rewritten as

$$\boldsymbol{\nabla}\cdot\left(u_0^2\boldsymbol{\nabla}\psi\right) = 0 . \tag{3.14}$$

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. (2.138), (2.139), (2.140) and (2.141)

$$\nabla \times \mathbf{H} = -ik_0\epsilon\mathbf{E} , \tag{3.15}$$

$$\nabla \times \mathbf{E} = ik_0\mu\mathbf{H} , \tag{3.16}$$

$$\nabla \cdot (\epsilon\mathbf{E}) = 0 , \tag{3.17}$$

$$\nabla \cdot (\mu\mathbf{H}) = 0 , \tag{3.18}$$

where

$$k_0 = \frac{\omega}{c} . \tag{3.19}$$

Recall that Eqs. (2.138), (2.139), (2.140) and (2.141) have been derived based on the assumptions that the medium is inhomogeneous, free of sources, isotropic, linear and stationary, the conductivity σ 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

$$\mathbf{E}(\mathbf{r}) = \exp[ik_0\psi(\mathbf{r})] \sum_{m=0}^{\infty} \frac{\mathbf{E}_m(\mathbf{r})}{(ik_0)^m} , \tag{3.20}$$

and

$$\mathbf{H}(\mathbf{r}) = \exp[ik_0\psi(\mathbf{r})] \sum_{m=0}^{\infty} \frac{\mathbf{H}_m(\mathbf{r})}{(ik_0)^m} . \tag{3.21}$$

Exercise 3.2.1. Show that

$$\nabla \times \mathbf{H}_m = -\epsilon\mathbf{E}_{m+1} - (\nabla\psi) \times \mathbf{H}_{m+1} , \tag{3.22}$$

$$\nabla \times \mathbf{E}_m = \mu\mathbf{H}_{m+1} - (\nabla\psi) \times \mathbf{E}_{m+1} , \tag{3.23}$$

$$\nabla \cdot (\epsilon\mathbf{E}_m) = -\epsilon\mathbf{E}_{m+1} \cdot \nabla\psi , \tag{3.24}$$

$$\nabla \cdot (\mu\mathbf{H}_m) = -\mu\mathbf{H}_{m+1} \cdot \nabla\psi . \tag{3.25}$$

Solution 3.2.1. Substituting Eqs. (3.20) and (3.21) into Eqs. (2.138), (2.139), (2.140) and (2.141) leads to

$$\sum_{m=0}^{\infty} \frac{1}{(ik_0)^m} \nabla \times [\mathbf{H}_m \exp(ik_0\psi)] = -\epsilon \exp(ik_0\psi) \sum_{m=-1}^{\infty} \frac{1}{(ik_0)^m} \mathbf{E}_{m+1} , \tag{3.26}$$

$$\sum_{m=0}^{\infty} \frac{1}{(ik_0)^m} \nabla \times [\mathbf{E}_m \exp(ik_0\psi)] = \mu \exp(ik_0\psi) \sum_{m=-1}^{\infty} \frac{1}{(ik_0)^m} \mathbf{H}_{m+1} , \tag{3.27}$$

$$\sum_{m=0}^{\infty} \frac{1}{(ik_0)^m} \nabla \cdot [\epsilon \mathbf{E}_m \exp(ik_0\psi)] = 0 , \tag{3.28}$$

and

$$\sum_{m=0}^{\infty} \frac{1}{(ik_0)^m} \nabla \cdot [\mu \mathbf{H}_m \exp(ik_0\psi)] = 0 . \tag{3.29}$$

By using the vectors identities (2.149) and (2.150) one obtains

$$\sum_{m=0}^{\infty} \frac{1}{(ik_0)^m} \nabla \times \mathbf{H}_m = \sum_{m=-1}^{\infty} \frac{1}{(ik_0)^m} [-\epsilon \mathbf{E}_{m+1} - (\nabla\psi) \times \mathbf{H}_{m+1}] , \tag{3.30}$$

$$\sum_{m=0}^{\infty} \frac{1}{(ik_0)^m} \nabla \times \mathbf{E}_m = \sum_{m=-1}^{\infty} \frac{1}{(ik_0)^m} [\mu \mathbf{H}_{m+1} - (\nabla\psi) \times \mathbf{E}_{m+1}] , \tag{3.31}$$

$$\sum_{m=0}^{\infty} \frac{1}{(ik_0)^m} \nabla \cdot (\epsilon \mathbf{E}_m) + \sum_{m=-1}^{\infty} \frac{1}{(ik_0)^m} \epsilon \mathbf{E}_{m+1} \cdot \nabla\psi = 0 , \tag{3.32}$$

$$\sum_{m=0}^{\infty} \frac{1}{(ik_0)^m} \nabla \cdot (\mu \mathbf{H}_m) + \sum_{m=-1}^{\infty} \frac{1}{(ik_0)^m} \mu \mathbf{H}_{m+1} \cdot \nabla\psi = 0 , \tag{3.33}$$

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 (2.138), (2.139), (2.140) and (2.141) while keeping only k_0 independent terms one obtains

$$\epsilon \mathbf{E}_0 + (\nabla\psi) \times \mathbf{H}_0 = 0 , \tag{3.34}$$

$$\mu \mathbf{H}_0 - (\nabla\psi) \times \mathbf{E}_0 = 0 , \tag{3.35}$$

$$\mathbf{E}_0 \cdot \nabla\psi = 0 , \tag{3.36}$$

$$\mathbf{H}_0 \cdot \nabla\psi = 0 . \tag{3.37}$$

Equations (3.34) and (3.35) imply that

$$n^2 \mathbf{E}_0 + (\nabla\psi) \times [(\nabla\psi) \times \mathbf{E}_0] = 0 , \tag{3.38}$$

thus, by using the vector identity (1.96), which is given by

$$\mathbf{A} \times (\mathbf{B} \times \mathbf{C}) = (\mathbf{A} \cdot \mathbf{C}) \mathbf{B} - (\mathbf{A} \cdot \mathbf{B}) \mathbf{C} , \tag{3.39}$$

together with Eq. (3.36) one obtains

$$\left[n^2 - (\nabla\psi)^2 \right] \mathbf{E}_0 = 0 . \tag{3.40}$$

The assumption that $\mathbf{E}_0$ does not vanish everywhere leads to the eikonal equation [see Eq. (3.7)]

$$(\nabla\psi)^2 = n^2 . \tag{3.41}$$

3.3 Optical Rays

The optical rays are defined as trajectories orthogonal to the wave front surfaces of constant ψ. Let $\mathbf{r}(s)$ be an optical ray with arc-length parametrization, namely $\mathrm{d}\mathbf{r}/\mathrm{d}s = \hat{\mathbf{s}}$, where $\hat{\mathbf{s}}$ is a unit vector. The ray equation reads

$$\frac{\mathrm{d}\mathbf{r}}{\mathrm{d}s} = \frac{\boldsymbol{\nabla}\psi}{n} = \hat{\mathbf{s}} . \tag{3.42}$$

The normal unit vector $\hat{\boldsymbol{\nu}}$ and the curvature κ are defined by the relation

$$\frac{\mathrm{d}\hat{\mathbf{s}}}{\mathrm{d}s} = \kappa\hat{\boldsymbol{\nu}} . \tag{3.43}$$

One can easily show that $\hat{\boldsymbol{\nu}} \cdot \hat{\mathbf{s}} = 0$ by taking the derivative of the relation $\hat{\mathbf{s}} \cdot \hat{\mathbf{s}} = 1$ with respect to s. The vectors $\hat{\mathbf{s}}$, $\hat{\boldsymbol{\nu}}$ together with the binormal unit vector $\hat{\mathbf{b}}$, which is defined by $\hat{\mathbf{b}} = \hat{\mathbf{s}} \times \hat{\boldsymbol{\nu}}$, form a local orthonormal coordinate frame.

Claim. The following holds

$$\frac{\mathrm{d}}{\mathrm{d}s}\begin{pmatrix} \hat{\mathbf{s}} \\ \hat{\boldsymbol{\nu}} \\ \hat{\mathbf{b}} \end{pmatrix} = \begin{pmatrix} 0 & \kappa & 0 \\ -\kappa & 0 & \tau \\ 0 & -\tau & 0 \end{pmatrix}\begin{pmatrix} \hat{\mathbf{s}} \\ \hat{\boldsymbol{\nu}} \\ \hat{\mathbf{b}} \end{pmatrix} , \tag{3.44}$$

where τ is real.

Proof. By taking the derivative of $\hat{\mathbf{s}} \cdot \hat{\boldsymbol{\nu}} = 0$ with respect to s one finds that

$$\hat{\mathbf{s}} \cdot \frac{\mathrm{d}\hat{\boldsymbol{\nu}}}{\mathrm{d}s} = -\kappa . \tag{3.45}$$

Similarly, by taking the derivative of $\hat{\mathbf{b}} \cdot \hat{\boldsymbol{\nu}} = 0$ with respect to s one obtains

$$\hat{\mathbf{b}} \cdot \frac{\mathrm{d}\hat{\boldsymbol{\nu}}}{\mathrm{d}s} = -\hat{\boldsymbol{\nu}} \cdot \frac{\mathrm{d}\hat{\mathbf{b}}}{\mathrm{d}s} . \tag{3.46}$$

By employing the definition $\hat{\mathbf{b}} = \hat{\mathbf{s}} \times \hat{\boldsymbol{\nu}}$ and Eq. (3.43) one finds that

$$\frac{\mathrm{d}\hat{\mathbf{b}}}{\mathrm{d}s} = \hat{\mathbf{s}} \times \frac{\mathrm{d}\hat{\boldsymbol{\nu}}}{\mathrm{d}s} , \tag{3.47}$$

thus

$$\hat{\mathbf{s}} \cdot \frac{\mathrm{d}\hat{\mathbf{b}}}{\mathrm{d}s} = 0 . \tag{3.48}$$

Moreover, by taking the derivative of $\hat{\mathbf{b}} \cdot \hat{\mathbf{b}} = 1$ with respect to s one finds that

$$\hat{\mathbf{b}} \cdot \frac{\mathrm{d}\hat{\mathbf{b}}}{\mathrm{d}s} = 0 \; , \tag{3.49}$$

thus $\mathrm{d}\hat{\mathbf{b}}/\mathrm{d}s$ is parallel to $\hat{\boldsymbol{\nu}}$. The torsion τ is defined as

$$\frac{\mathrm{d}\hat{\mathbf{b}}}{\mathrm{d}s} = -\tau\hat{\boldsymbol{\nu}} \; . \tag{3.50}$$

The above results (and definitions) can be summarized by Eq. (3.44).

Alternatively, Eq. (3.44) can be expressed as

$$\frac{\mathrm{d}}{\mathrm{d}s}\begin{pmatrix} \hat{\mathbf{s}} \\ \hat{\boldsymbol{\nu}} \\ \hat{\mathbf{b}} \end{pmatrix} = \boldsymbol{\delta} \times \begin{pmatrix} \hat{\mathbf{s}} \\ \hat{\boldsymbol{\nu}} \\ \hat{\mathbf{b}} \end{pmatrix} \; , \tag{3.51}$$

where

$$\boldsymbol{\delta} = \tau\hat{\mathbf{s}} + \kappa\hat{\mathbf{b}} \; . \tag{3.52}$$

Exercise 3.3.1. Show that for any closed curve C the following holds

$$\oint_C (n\hat{\mathbf{s}}) \cdot \mathrm{d}\mathbf{l} = 0 \; . \tag{3.53}$$

Solution 3.3.1. With the help of Eq. (3.42) one finds that

$$\boldsymbol{\nabla} \times (n\hat{\mathbf{s}}) = \boldsymbol{\nabla} \times \boldsymbol{\nabla}\psi = 0 \; , \tag{3.54}$$

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 θ_{i}, θ_{r} and θ_{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 θ_{i}, θ_{r} and θ_{t}. By employing the Lagrange's integral invariant (3.53), which implies that the tangential component of $n\hat{\mathbf{s}}$ is continuous [compare with Eq. (2.78)], one obtains the relation

$$\theta_{\mathrm{i}} = \theta_{\mathrm{r}} \; ,$$

and the so-called Snell's law, which is given by [compare with Eq. (2.237)]

$$n_1 \sin\theta_{\mathrm{i}} = n_2 \sin\theta_{\mathrm{t}} \; . \tag{3.55}$$

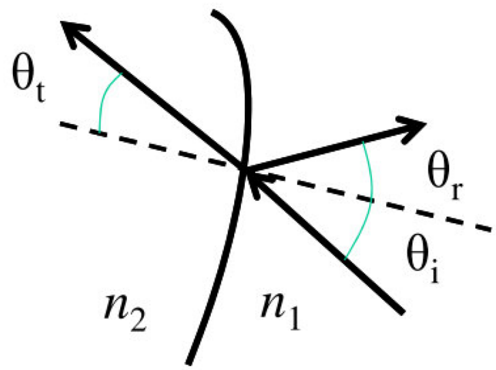

Fig. 3.1. Incident, reflected and refracted rays.

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

$$\kappa\hat{\boldsymbol{\nu}} = \boldsymbol{\nabla}(\log n) - \hat{\mathbf{s}}\frac{\mathrm{d}(\log n)}{\mathrm{d}s} . \tag{3.56}$$

Proof. Taking the derivative $\mathrm{d}/\mathrm{d}s$ of Eq. (3.42) leads to

$$\frac{\mathrm{d}}{\mathrm{d}s}\left(n\frac{\mathrm{d}\mathbf{r}}{\mathrm{d}s}\right) = \frac{\mathrm{d}}{\mathrm{d}s}\boldsymbol{\nabla}\psi = \frac{1}{n}(\boldsymbol{\nabla}\psi\cdot\boldsymbol{\nabla})\boldsymbol{\nabla}\psi . \tag{3.57}$$

By using the vector identities

$$\boldsymbol{\nabla}\left(\mathbf{A}^2\right) = 2\left[\mathbf{A}\times(\boldsymbol{\nabla}\times\mathbf{A}) + (\mathbf{A}\cdot\boldsymbol{\nabla})\mathbf{A}\right] , \tag{3.58}$$

and

$$\boldsymbol{\nabla}\times(\boldsymbol{\nabla}\psi) = 0 , \tag{3.59}$$

one finds that

$$\frac{\mathrm{d}}{\mathrm{d}s}\left(n\frac{\mathrm{d}\mathbf{r}}{\mathrm{d}s}\right) = \frac{1}{2n}\boldsymbol{\nabla}(\boldsymbol{\nabla}\psi)^2 , \tag{3.60}$$

or [see Eq. (3.41)]

$$\frac{\mathrm{d}}{\mathrm{d}s}\left(n\frac{\mathrm{d}\mathbf{r}}{\mathrm{d}s}\right) = \frac{1}{2n}\boldsymbol{\nabla}n^2 = \boldsymbol{\nabla}n . \tag{3.61}$$

Moreover, the following holds

$$\frac{\mathrm{d}}{\mathrm{d}s}\left(n\frac{\mathrm{d}\mathbf{r}}{\mathrm{d}s}\right)=\hat{\mathbf{s}}\frac{\mathrm{d}n}{\mathrm{d}s}+n\kappa\hat{\boldsymbol{\nu}}\ . \tag{3.62}$$

The last two results directly lead to Eq. (3.56).

By multiplying Eq. (3.56) by $\hat{\boldsymbol{\nu}}$ one obtains

$$\kappa=\hat{\boldsymbol{\nu}}\cdot\boldsymbol{\nabla}\left(\log n\right)\ , \tag{3.63}$$

thus the ray is curved towards the regime of higher n. One also finds that the vector $\boldsymbol{\nabla}n$ is in the so-called osculating plane of the ray ($\hat{\mathbf{s}}\hat{\boldsymbol{\nu}}$ plane).

Exercise 3.3.2. Show that

$$\frac{\mathrm{d}\hat{\mathbf{s}}}{\mathrm{d}s}=\hat{\mathbf{s}}\times\left(\boldsymbol{\nabla}\left(\log n\right)\times\hat{\mathbf{s}}\right)\ . \tag{3.64}$$

Solution 3.3.2. The general identity [see Eq. (1.96)]

$$\mathbf{A}\times\left(\mathbf{B}\times\mathbf{A}\right)=\left(\mathbf{A}\cdot\mathbf{A}\right)\mathbf{B}-\left(\mathbf{A}\cdot\mathbf{B}\right)\mathbf{A}\ , \tag{3.65}$$

together with Eq. (3.56) lead to Eq. (3.64).

Exercise 3.3.3. Consider a parametrization of an optical ray given by

$$\mathbf{r}\left(s\right)=\mathbf{r}\left(\sigma\left(s\right)\right)\ , \tag{3.66}$$

where

$$\frac{\mathrm{d}s}{\mathrm{d}\sigma}=n\ . \tag{3.67}$$

Show that

$$\frac{\mathrm{d}^2\mathbf{r}}{\mathrm{d}\sigma^2}=-\boldsymbol{\nabla}U\ , \tag{3.68}$$

where

$$U=-\frac{n^2}{2}\ . \tag{3.69}$$

Solution 3.3.3. With the help of Eqs. (3.61) and the relation [see (3.67)]

$$\frac{\mathrm{d}}{\mathrm{d}s}=\frac{1}{n}\frac{\mathrm{d}}{\mathrm{d}\sigma}\ , \tag{3.70}$$

one finds that

$$\frac{\mathrm{d}^2\mathbf{r}}{\mathrm{d}\sigma^2}=\boldsymbol{\nabla}\left(\frac{n^2}{2}\right)\ . \tag{3.71}$$

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 $n\mathbf{r}\times\hat{\mathbf{s}}$ is a constant along an optical ray traveling in that medium.

Solution 3.3.4. The following holds

$$\frac{\mathrm{d}}{\mathrm{d}s}\left(n\mathbf{r}\times\hat{\mathbf{s}}\right) = \frac{\mathrm{d}\mathbf{r}}{\mathrm{d}s}\times\left(n\hat{\mathbf{s}}\right) + \mathbf{r}\times\frac{\mathrm{d}\left(n\hat{\mathbf{s}}\right)}{\mathrm{d}s} \ . \tag{3.73}$$

The first term on the right hand side of Eq. (3.73) vanishes since $\mathrm{d}\mathbf{r}/\mathrm{d}s = \hat{\mathbf{s}}$ [see Eq. (3.42)]. The same identity (3.42) together with Eq. (3.61) lead to

$$\frac{\mathrm{d}\left(n\hat{\mathbf{s}}\right)}{\mathrm{d}s} = \boldsymbol{\nabla} n \ , \tag{3.74}$$

and thus also the second term on the right hand side of Eq. (3.73) vanishes provided that $n = n(r)$, which implies that

$$\frac{\mathrm{d}}{\mathrm{d}s}\left(n\mathbf{r}\times\hat{\mathbf{s}}\right) = 0 \ , \tag{3.75}$$

i.e. $n\mathbf{r}\times\hat{\mathbf{s}}$ 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. (2.151)]. 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 $\mathbf{x}_{\mathrm{C}}(q)$

$$\mathbf{x}_{\mathrm{C}}(q) = \left(x_{\mathrm{C}1}(q), x_{\mathrm{C}2}(q), x_{\mathrm{C}3}(q)\right) \ , \tag{3.76}$$

where $q \in [q_1, q_2]$. The time of flight T of light traveling along the curve C is given by

$$T = c^{-1}\int_{q_1}^{q_2} \mathrm{d}q\, n\left(\mathbf{x}_{\mathrm{C}}\right)\left|\dot{\mathbf{x}}_{\mathrm{C}}\right| \ , \tag{3.77}$$

where

$$\dot{\mathbf{x}}_{\mathrm{C}} = \frac{\mathrm{d}\mathbf{x}_{\mathrm{C}}}{\mathrm{d}q} \ , \tag{3.78}$$

or, alternatively by

$$T = c^{-1} \int_{q_1}^{q_2} \mathrm{d}q \, \mathcal{L}\left(\mathbf{x}_{\mathrm{C}}, \dot{\mathbf{x}}_{\mathrm{C}}\right) \ , \tag{3.79}$$

where $\mathcal{L}$ is given by

$$\mathcal{L}\left(\mathbf{x}_{\mathrm{C}}, \dot{\mathbf{x}}_{\mathrm{C}}\right) = n\left(\mathbf{x}_{\mathrm{C}}\right) \sqrt{\dot{\mathbf{x}}_{\mathrm{C}}^2} \ , \tag{3.80}$$

and

$$\dot{\mathbf{x}}_{\mathrm{C}} = \left(\dot{x}_{\mathrm{C}1}\left(q\right), \dot{x}_{\mathrm{C}2}\left(q\right), \dot{x}_{\mathrm{C}3}\left(q\right)\right) \ . \tag{3.81}$$

Consider another curve C', which is assumed to be infinitesimally close to the curve C, and which is given by

$$\mathbf{x}'_{\mathrm{C}}\left(q\right) = \mathbf{x}_{\mathrm{C}}\left(q\right) + \delta\mathbf{x}\left(q\right) \ . \tag{3.82}$$

It is assumed that $\delta\mathbf{x}\left(q_1\right) = \delta\mathbf{x}\left(q_2\right) = 0$, i.e. the curves C and C' are taken to have the same initial and final points. To lowest nonvanishing order in $\delta\mathbf{x}\left(q\right) = \left(\delta_{\mathrm{C}1}\left(q\right), \delta_{\mathrm{C}2}\left(q\right), \delta_{\mathrm{C}3}\left(q\right)\right)$ the change $\delta\mathcal{L}$ in $\mathcal{L}$ is given by

$$\delta\mathcal{L} = \sum_{n=1}^{3} \frac{\partial\mathcal{L}}{\partial x_{\mathrm{C}n}} \delta_{\mathrm{C}n} + \frac{\partial\mathcal{L}}{\partial \dot{x}_{\mathrm{C}n}} \frac{\mathrm{d}\left(\delta_{\mathrm{C}n}\right)}{\mathrm{d}q} \ , \tag{3.83}$$

thus, to lowest nonvanishing order in $\delta\mathbf{x}$ the change δT in the time of flight T is given by

$$\begin{aligned} \delta T &= c^{-1} \int_{q_1}^{q_2} \mathrm{d}q \, \delta\mathcal{L}\left(\mathbf{x}_{\mathrm{C}}, \dot{\mathbf{x}}_{\mathrm{C}}\right) \\ &= c^{-1} \int_{q_1}^{q_2} \mathrm{d}q \, \sum_{n=1}^{3} \frac{\partial\mathcal{L}}{\partial x_{\mathrm{C}n}} \delta_{\mathrm{C}n} + \frac{\partial\mathcal{L}}{\partial \dot{x}_{\mathrm{C}n}} \frac{\mathrm{d}\left(\delta_{\mathrm{C}n}\right)}{\mathrm{d}q} \ . \end{aligned} \tag{3.84}$$

Integration by parts yields [recall that $\delta\mathbf{x}\left(q_1\right) = \delta\mathbf{x}\left(q_2\right) = 0$]

$$\delta T = c^{-1} \int_{q_1}^{q_2} \mathrm{d}q \, \sum_{n=1}^{3} \left(\frac{\partial\mathcal{L}}{\partial x_{\mathrm{C}n}} - \frac{\mathrm{d}}{\mathrm{d}q} \frac{\partial\mathcal{L}}{\partial \dot{x}_{\mathrm{C}n}} \right) \delta_{\mathrm{C}n} \ , \tag{3.85}$$

or

$$\delta T = c^{-1} \int_{q_1}^{q_2} \mathrm{d}q \, \left(\sqrt{\dot{\mathbf{x}}_{\mathrm{C}}^2} \boldsymbol{\nabla} n - \frac{\mathrm{d}}{\mathrm{d}q} \frac{n\dot{\mathbf{x}}_{\mathrm{C}}}{\sqrt{\dot{\mathbf{x}}_{\mathrm{C}}^2}} \right) \cdot \delta\mathbf{x} \ . \tag{3.86}$$

Claim. The requirement the $\delta T = 0$ for an arbitrary $\delta\mathbf{x}$ implies that the curve C satisfies Eq. (3.61).

Proof. The requirement implies that [see Eq. (3.86)]

$$\frac{\mathrm{d}}{\mathrm{d}q}\frac{n\dot{\mathbf{x}}_{\mathrm{C}}}{\sqrt{\dot{\mathbf{x}}_{\mathrm{C}}^2}} = \sqrt{\dot{\mathbf{x}}_{\mathrm{C}}^2}\boldsymbol{\nabla} n \; . \tag{3.87}$$

With arc-length parameterization the curve C is expressed as $\mathbf{x}_{\mathrm{C}}(q(s))$, where

$$\left|\frac{\mathrm{d}\mathbf{x}_{\mathrm{C}}}{\mathrm{d}s}\right| = 1 \; . \tag{3.88}$$

With the help of the following relation

$$\frac{\mathrm{d}}{\mathrm{d}q} = \frac{\mathrm{d}s}{\mathrm{d}q}\frac{\mathrm{d}}{\mathrm{d}s} \; , \tag{3.89}$$

Eq. (3.87) becomes

$$\frac{\mathrm{d}}{\mathrm{d}s}\left(n\frac{\mathrm{d}\mathbf{x}_C}{\mathrm{d}s}\right) = \boldsymbol{\nabla} n \; , \tag{3.90}$$

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 $\mathbf{r}_1$ to a final one $\mathbf{r}_2$ the time of flight T is given by

$$T = \frac{\psi(\mathbf{r}_2) - \psi(\mathbf{r}_1)}{c} \; , \tag{3.91}$$

where ψ is the eikonal.

Solution 3.3.5. By multiplying Eq. (3.42) by the unit vector $\mathrm{d}\mathbf{r}/\mathrm{d}s = \hat{\mathbf{s}}$ one obtains

$$n = \frac{\mathrm{d}\mathbf{r}}{\mathrm{d}s}\cdot\boldsymbol{\nabla}\psi \; , \tag{3.92}$$

thus the time of flight T for an optical ray $\mathbf{r}(s)$ in arc-length parametrization is given by [see Eq. (3.77)]

$$\begin{aligned} T &= c^{-1}\int_{s_1}^{s_2}\mathrm{d}s\; n\left|\frac{\mathrm{d}\mathbf{r}}{\mathrm{d}s}\right| \\ &= c^{-1}\int_{s_1}^{s_2}\mathrm{d}s\; n \\ &= c^{-1}\int_{s_1}^{s_2}\mathrm{d}s\; \left(\frac{\mathrm{d}\mathbf{r}}{\mathrm{d}s}\cdot\boldsymbol{\nabla}\psi\right) \\ &= c^{-1}\int_{\mathbf{r}_1}^{\mathbf{r}_2}\mathrm{d}\mathbf{r}\cdot\boldsymbol{\nabla}\psi \; , \end{aligned} \tag{3.93}$$

and therefore Eq. (3.91) holds.

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, ϕ, 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)]

$$\frac{\partial \mathcal{L}}{\partial x_{\mathrm{C}n}} = \frac{\mathrm{d}}{\mathrm{d}q}\frac{\partial \mathcal{L}}{\partial \dot{x}_{\mathrm{C}n}} , \tag{3.94}$$

where $\mathcal{L}$ is given by Eq. (3.80). For the case of cylindrical coordinates one obtains

$$\frac{\partial \mathcal{L}}{\partial r} = \frac{\mathrm{d}}{\mathrm{d}q}\frac{\partial \mathcal{L}}{\partial \dot{r}} , \tag{3.95}$$

$$\frac{\partial \mathcal{L}}{\partial \phi} = \frac{\mathrm{d}}{\mathrm{d}q}\frac{\partial \mathcal{L}}{\partial \dot{\phi}} , \tag{3.96}$$

$$\frac{\partial \mathcal{L}}{\partial z} = \frac{\mathrm{d}}{\mathrm{d}q}\frac{\partial \mathcal{L}}{\partial \dot{z}} . \tag{3.97}$$

The following holds

$$\mathbf{x}_{\mathrm{C}} = (r\cos\phi, r\sin\phi, z) , \tag{3.98}$$

$$\dot{\mathbf{x}}_{\mathrm{C}} = \left(\frac{\mathrm{d}r}{\mathrm{d}q}\cos\phi - r\sin\phi\frac{\mathrm{d}\phi}{\mathrm{d}q}, \frac{\mathrm{d}r}{\mathrm{d}q}\sin\phi + r\cos\phi\frac{\mathrm{d}\phi}{\mathrm{d}q}, \frac{\mathrm{d}z}{\mathrm{d}q}\right) , \tag{3.99}$$

and thus

$$\dot{\mathbf{x}}_{\mathrm{C}}^2 = \left(\frac{\mathrm{d}r}{\mathrm{d}q}\right)^2 + \left(r\frac{\mathrm{d}\phi}{\mathrm{d}q}\right)^2 + \left(\frac{\mathrm{d}z}{\mathrm{d}q}\right)^2 , \tag{3.100}$$

and [see Eq. (3.80)]

$$\mathcal{L} = n\sqrt{\dot{\mathbf{x}}_{\mathrm{C}}^2} = n\sqrt{\dot{r}^2 + \left(r\dot{\phi}\right)^2 + \dot{z}^2} . \tag{3.101}$$

With the help of the above result Eqs. (3.95). (3.96) and (3.97) become

$$\sqrt{\dot{\mathbf{x}}_{\mathrm{C}}^2}\frac{\partial n}{\partial r} + n\frac{\partial\sqrt{\dot{\mathbf{x}}_{\mathrm{C}}^2}}{\partial r} = \frac{\mathrm{d}}{\mathrm{d}q}\frac{n\dot{r}}{\sqrt{\dot{\mathbf{x}}_{\mathrm{C}}^2}} , \tag{3.102}$$

$$\sqrt{\dot{\mathbf{x}}_{\mathrm{C}}^2}\frac{\partial n}{\partial \phi} = \frac{\mathrm{d}}{\mathrm{d}q}\frac{nr^2\dot{\phi}}{\sqrt{\dot{\mathbf{x}}_{\mathrm{C}}^2}} , \tag{3.103}$$

$$\sqrt{\dot{\mathbf{x}}_{\mathrm{C}}^2}\frac{\partial n}{\partial z} = \frac{\mathrm{d}}{\mathrm{d}q}\frac{n\dot{z}}{\sqrt{\dot{\mathbf{x}}_{\mathrm{C}}^2}} . \tag{3.104}$$

With arc-length parameterization the curve C is expressed as $\mathbf{x}_{\mathrm{C}}(q(s))$, where

$$\left|\frac{\mathrm{d}\mathbf{x}_{\mathrm{C}}}{\mathrm{d}s}\right| = 1 \,, \tag{3.105}$$

and where

$$\frac{\mathrm{d}}{\mathrm{d}q} = \frac{\mathrm{d}s}{\mathrm{d}q}\frac{\mathrm{d}}{\mathrm{d}s} \,, \tag{3.106}$$

and thus the following holds

$$\sqrt{\dot{\mathbf{x}}_{\mathrm{C}}^2} = \frac{\mathrm{d}s}{\mathrm{d}q} \,, \tag{3.107}$$

and Eqs. (3.102), (3.103) and (3.104) become

$$\frac{\partial n}{\partial r} + nr\left(\frac{\mathrm{d}\phi}{\mathrm{d}s}\right)^2 = \frac{\mathrm{d}}{\mathrm{d}s}\left(n\frac{\mathrm{d}r}{\mathrm{d}s}\right) \,, \tag{3.108}$$

$$\frac{\partial n}{\partial \phi} = \frac{\mathrm{d}}{\mathrm{d}s}\left(nr^2\frac{\mathrm{d}\phi}{\mathrm{d}s}\right) \,, \tag{3.109}$$

$$\frac{\partial n}{\partial z} = \frac{\mathrm{d}}{\mathrm{d}s}\left(n\frac{\mathrm{d}z}{\mathrm{d}s}\right) \,. \tag{3.110}$$

For the case of cylindrically symmetric medium $n = n(r)$, and thus

$$\frac{\partial n}{\partial r} + nr\left(\frac{\mathrm{d}\phi}{\mathrm{d}s}\right)^2 = \frac{\mathrm{d}}{\mathrm{d}s}\left(n\frac{\mathrm{d}r}{\mathrm{d}s}\right) \,, \tag{3.111}$$

$$0 = \frac{\mathrm{d}}{\mathrm{d}s}\left(nr^2\frac{\mathrm{d}\phi}{\mathrm{d}s}\right) \,, \tag{3.112}$$

$$0 = \frac{\mathrm{d}}{\mathrm{d}s}\left(n\frac{\mathrm{d}z}{\mathrm{d}s}\right) \,. \tag{3.113}$$

3.4 Transport Equation

The so-called transport equation [see Eq. (3.114) below] is needed in order to evaluate the evolution of polarization along an optical ray.

Claim. The vector $\mathbf{E}_0$ satisfies the transport equation, which is given by

$$2(\boldsymbol{\nabla}\psi\cdot\boldsymbol{\nabla})\,\mathbf{E}_0 + \mathbf{E}_0\left[\boldsymbol{\nabla}^2\psi - \boldsymbol{\nabla}(\log\mu)\cdot\boldsymbol{\nabla}\psi\right] + 2\left[\mathbf{E}_0\cdot\boldsymbol{\nabla}(\log n)\right]\boldsymbol{\nabla}\psi = 0 \,. \tag{3.114}$$

Proof. Eliminating $\mathbf{H}_{m+1}$ and $\mathbf{H}_m$ [lowering the index in Eq. (3.23) by one] from Eq. (3.23) and substituting into Eq. (3.22) yield

$$\begin{aligned}&\boldsymbol{\nabla}\times\left[\frac{1}{\mu}\boldsymbol{\nabla}\times\mathbf{E}_{m-1}\right]+\boldsymbol{\nabla}\times\left[\frac{\boldsymbol{\nabla}\psi}{\mu}\times\mathbf{E}_m\right]\\&=-\epsilon\mathbf{E}_{m+1}-\frac{\boldsymbol{\nabla}\psi}{\mu}\times\left(\boldsymbol{\nabla}\times\mathbf{E}_m\right)-\left(\boldsymbol{\nabla}\psi\right)\times\left[\frac{1}{\mu}\left(\boldsymbol{\nabla}\psi\right)\times\mathbf{E}_{m+1}\right]\ .\end{aligned}\tag{3.115}$$

The two terms involved with $\mathbf{E}_m$ are treated as follows. Using the vector identities

$$\boldsymbol{\nabla}\left(\mathbf{A}\cdot\mathbf{B}\right)=\mathbf{A}\times\left(\boldsymbol{\nabla}\times\mathbf{B}\right)+\mathbf{B}\times\left(\boldsymbol{\nabla}\times\mathbf{A}\right)+\left(\mathbf{B}\cdot\boldsymbol{\nabla}\right)\mathbf{A}+\left(\mathbf{A}\cdot\boldsymbol{\nabla}\right)\mathbf{B}\ ,\tag{3.116}$$

and (1.128), which is given by

$$\boldsymbol{\nabla}\times\left(\mathbf{A}\times\mathbf{B}\right)=\mathbf{A}\left(\boldsymbol{\nabla}\cdot\mathbf{B}\right)-\mathbf{B}\left(\boldsymbol{\nabla}\cdot\mathbf{A}\right)+\left(\mathbf{B}\cdot\boldsymbol{\nabla}\right)\mathbf{A}-\left(\mathbf{A}\cdot\boldsymbol{\nabla}\right)\mathbf{B}\ ,\tag{3.117}$$

one finds that

$$\begin{aligned}&\boldsymbol{\nabla}\times\left(\mathbf{A}\times\mathbf{B}\right)+\mathbf{A}\times\left(\boldsymbol{\nabla}\times\mathbf{B}\right)\\&=\mathbf{A}\left(\boldsymbol{\nabla}\cdot\mathbf{B}\right)-\mathbf{B}\left(\boldsymbol{\nabla}\cdot\mathbf{A}\right)+\boldsymbol{\nabla}\left(\mathbf{A}\cdot\mathbf{B}\right)-\mathbf{B}\times\left(\boldsymbol{\nabla}\times\mathbf{A}\right)-2\left(\mathbf{A}\cdot\boldsymbol{\nabla}\right)\mathbf{B}\ ,\end{aligned}\tag{3.118}$$

thus

$$\begin{aligned}&\boldsymbol{\nabla}\times\left(\frac{\boldsymbol{\nabla}\psi}{\mu}\times\mathbf{E}_m\right)+\frac{\boldsymbol{\nabla}\psi}{\mu}\times\left(\boldsymbol{\nabla}\times\mathbf{E}_m\right)\\&=\frac{\boldsymbol{\nabla}\psi}{\mu}\left(\boldsymbol{\nabla}\cdot\mathbf{E}_m\right)-\mathbf{E}_m\left(\boldsymbol{\nabla}\cdot\frac{\boldsymbol{\nabla}\psi}{\mu}\right)+\boldsymbol{\nabla}\left(\frac{\boldsymbol{\nabla}\psi}{\mu}\cdot\mathbf{E}_m\right)-\mathbf{E}_m\times\left(\boldsymbol{\nabla}\times\frac{\boldsymbol{\nabla}\psi}{\mu}\right)-2\left(\frac{\boldsymbol{\nabla}\psi}{\mu}\cdot\boldsymbol{\nabla}\right)\mathbf{E}_m\ .\end{aligned}\tag{3.119}$$

By using Eq. (2.150) and the vector identity $\boldsymbol{\nabla}\times\left(\boldsymbol{\nabla}f\right)=0$ one obtains

$$\boldsymbol{\nabla}\times\frac{\boldsymbol{\nabla}\psi}{\mu}=\frac{1}{\mu}\boldsymbol{\nabla}\times\boldsymbol{\nabla}\psi+\boldsymbol{\nabla}\left(\frac{1}{\mu}\right)\times\boldsymbol{\nabla}\psi=\boldsymbol{\nabla}\left(\frac{1}{\mu}\right)\times\boldsymbol{\nabla}\psi\ .\tag{3.120}$$

Substituting into Eq. (3.115) yields

$$\begin{aligned}&2\left(\boldsymbol{\nabla}\psi\cdot\boldsymbol{\nabla}\right)\mathbf{E}_m-\boldsymbol{\nabla}\psi\left(\boldsymbol{\nabla}\cdot\mathbf{E}_m\right)+\mu\mathbf{E}_m\left(\boldsymbol{\nabla}\cdot\frac{\boldsymbol{\nabla}\psi}{\mu}\right)\\&-\mu\boldsymbol{\nabla}\left(\frac{\boldsymbol{\nabla}\psi}{\mu}\cdot\mathbf{E}_m\right)+\mu\mathbf{E}_m\times\left(\boldsymbol{\nabla}\left(\frac{1}{\mu}\right)\times\boldsymbol{\nabla}\psi\right)\\&=n^2\mathbf{E}_{m+1}+\left(\boldsymbol{\nabla}\psi\right)\times\left[\left(\boldsymbol{\nabla}\psi\right)\times\mathbf{E}_{m+1}\right]+\mu\boldsymbol{\nabla}\times\left[\frac{1}{\mu}\boldsymbol{\nabla}\times\mathbf{E}_{m-1}\right]\ ,\end{aligned}\tag{3.121}$$

where $n = \sqrt{\epsilon\mu}$. By using Eq. (1.96) for the fifth term on the left hand side and for the second term on the right hand side one finds that

$$\begin{aligned}
&2\left(\boldsymbol{\nabla}\psi\cdot\boldsymbol{\nabla}\right)\mathbf{E}_m-\boldsymbol{\nabla}\psi\left(\boldsymbol{\nabla}\cdot\mathbf{E}_m\right)+\mu\mathbf{E}_m\left(\boldsymbol{\nabla}\cdot\frac{\boldsymbol{\nabla}\psi}{\mu}\right)\\
&-\mu\boldsymbol{\nabla}\left(\frac{\boldsymbol{\nabla}\psi}{\mu}\cdot\mathbf{E}_m\right)+\mu\left(\mathbf{E}_m\cdot\boldsymbol{\nabla}\psi\right)\boldsymbol{\nabla}\left(\frac{1}{\mu}\right)-\mu\left(\mathbf{E}_m\cdot\boldsymbol{\nabla}\left(\frac{1}{\mu}\right)\right)\boldsymbol{\nabla}\psi\\
&=\left[n^2-\left(\boldsymbol{\nabla}\psi\right)^2\right]\mathbf{E}_{m+1}+\left[\left(\boldsymbol{\nabla}\psi\right)\cdot\mathbf{E}_{m+1}\right]\left(\boldsymbol{\nabla}\psi\right)+\mu\boldsymbol{\nabla}\times\left[\frac{1}{\mu}\boldsymbol{\nabla}\times\mathbf{E}_{m-1}\right]\ .
\end{aligned}\tag{3.122}$$

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

$$-\boldsymbol{\nabla}\psi\cdot\mathbf{E}_m=\frac{1}{\epsilon}\boldsymbol{\nabla}\cdot\left(\epsilon\mathbf{E}_{m-1}\right)\ ,\tag{3.123}$$

is employed to rewrite the forth term on the left hand side, thus

$$\begin{aligned}
\boldsymbol{\nabla}\left(\frac{\boldsymbol{\nabla}\psi}{\mu}\cdot\mathbf{E}_m\right)&=\boldsymbol{\nabla}\left(\frac{1}{\mu}\right)\left(\boldsymbol{\nabla}\psi\cdot\mathbf{E}_m\right)+\frac{1}{\mu}\boldsymbol{\nabla}\left(\boldsymbol{\nabla}\psi\cdot\mathbf{E}_m\right)\\
&=\boldsymbol{\nabla}\left(\frac{1}{\mu}\right)\left(\boldsymbol{\nabla}\psi\cdot\mathbf{E}_m\right)-\frac{1}{\mu}\boldsymbol{\nabla}\left[\frac{1}{\epsilon}\boldsymbol{\nabla}\cdot\left(\epsilon\mathbf{E}_{m-1}\right)\right]\ ,
\end{aligned}\tag{3.124}$$

hence

$$\begin{aligned}
&2\left(\boldsymbol{\nabla}\psi\cdot\boldsymbol{\nabla}\right)\mathbf{E}_m-\boldsymbol{\nabla}\psi\left(\boldsymbol{\nabla}\cdot\mathbf{E}_m\right)+\mu\mathbf{E}_m\left(\boldsymbol{\nabla}\cdot\frac{\boldsymbol{\nabla}\psi}{\mu}\right)-\mu\left(\mathbf{E}_m\cdot\boldsymbol{\nabla}\left(\frac{1}{\mu}\right)\right)\boldsymbol{\nabla}\psi\\
&=\left[\left(\boldsymbol{\nabla}\psi\right)\cdot\mathbf{E}_{m+1}\right]\left(\boldsymbol{\nabla}\psi\right)+\mu\boldsymbol{\nabla}\times\left[\frac{1}{\mu}\boldsymbol{\nabla}\times\mathbf{E}_{m-1}\right]-\boldsymbol{\nabla}\left[\frac{1}{\epsilon}\boldsymbol{\nabla}\cdot\left(\epsilon\mathbf{E}_{m-1}\right)\right]\ .
\end{aligned}\tag{3.125}$$

Using Eqs. (3.24) and (2.149) for the second term on the left hand side one finds that

$$\begin{aligned}
-\boldsymbol{\nabla}\psi\boldsymbol{\nabla}\cdot\mathbf{E}_m&=-\boldsymbol{\nabla}\psi\left[\frac{1}{\epsilon}\boldsymbol{\nabla}\cdot\left(\epsilon\mathbf{E}_m\right)-\frac{1}{\epsilon}\mathbf{E}_m\cdot\boldsymbol{\nabla}\epsilon\right]\\
&=\boldsymbol{\nabla}\psi\left[\boldsymbol{\nabla}\psi\cdot\mathbf{E}_{m+1}+\frac{1}{\epsilon}\mathbf{E}_m\cdot\boldsymbol{\nabla}\epsilon\right]\ ,
\end{aligned}\tag{3.126}$$

thus

$$\begin{aligned}
&2\left(\boldsymbol{\nabla}\psi\cdot\boldsymbol{\nabla}\right)\mathbf{E}_m+\boldsymbol{\nabla}\psi\left(\frac{1}{\epsilon}\mathbf{E}_m\cdot\boldsymbol{\nabla}\epsilon\right)+\mu\mathbf{E}_m\left(\boldsymbol{\nabla}\cdot\frac{\boldsymbol{\nabla}\psi}{\mu}\right)-\mu\left(\mathbf{E}_m\cdot\boldsymbol{\nabla}\left(\frac{1}{\mu}\right)\right)\boldsymbol{\nabla}\psi\\
&=\mu\boldsymbol{\nabla}\times\left[\frac{1}{\mu}\boldsymbol{\nabla}\times\mathbf{E}_{m-1}\right]-\boldsymbol{\nabla}\left[\frac{1}{\epsilon}\boldsymbol{\nabla}\cdot\left(\epsilon\mathbf{E}_{m-1}\right)\right]\ .
\end{aligned}\tag{3.127}$$

After arranging terms this becomes

$$2\left(\boldsymbol{\nabla}\psi\cdot\boldsymbol{\nabla}\right)\mathbf{E}_m+\mathbf{E}_m\left[\boldsymbol{\nabla}^2\psi-\boldsymbol{\nabla}\left(\log\mu\right)\cdot\boldsymbol{\nabla}\psi\right]+2\left[\mathbf{E}_m\cdot\boldsymbol{\nabla}\left(\log n\right)\right]\boldsymbol{\nabla}\psi$$
$$=\mu\boldsymbol{\nabla}\times\left[\frac{1}{\mu}\boldsymbol{\nabla}\times\mathbf{E}_{m-1}\right]-\boldsymbol{\nabla}\left[\frac{1}{\epsilon}\boldsymbol{\nabla}\cdot\left(\epsilon\mathbf{E}_{m-1}\right)\right] . \tag{3.128}$$

For the case $m=0$ Eq. (3.128) becomes Eq. (3.114).

3.4.1 Polarization Evolution

In the exercise below the transport equation (3.114) is rewritten in terms of the unit vector $\hat{\mathbf{e}}_0$, which points in the direction of $\mathbf{E}_0$.

Exercise 3.4.1. Show that

$$\frac{\mathrm{d}}{\mathrm{d}s}\hat{\mathbf{e}}_0=-\kappa\left(\hat{\mathbf{e}}_0\cdot\hat{\boldsymbol{\nu}}\right)\hat{\mathbf{s}} , \tag{3.129}$$

where $\hat{\mathbf{e}}_0$ is a unit vector in the direction of $\mathbf{E}_0$

$$\hat{\mathbf{e}}_0\equiv\frac{\mathbf{E}_0}{\sqrt{\mathbf{E}_0\cdot\mathbf{E}_0^*}} . \tag{3.130}$$

Solution 3.4.1. By multiplying the transport equation (3.114) by $\mathbf{E}_0^*$, using Eq. (3.36), and taking the real part of the resulting equation one obtains

$$\left[\boldsymbol{\nabla}^2\psi-\left[\boldsymbol{\nabla}\left(\log\mu\right)\cdot\left(\boldsymbol{\nabla}\psi\right)\right]\right]\left(\mathbf{E}_0\cdot\mathbf{E}_0^*\right)+\left(\boldsymbol{\nabla}\psi\cdot\boldsymbol{\nabla}\right)\left(\mathbf{E}_0\cdot\mathbf{E}_0^*\right)=0 . \tag{3.131}$$

Substituting $\mathbf{E}_0=\sqrt{\mathbf{E}_0\cdot\mathbf{E}_0^*}\hat{\mathbf{e}}_0$ [see Eq. (3.130)] into Eq. (3.114) leads to

$$\frac{1}{2}\left[\boldsymbol{\nabla}^2\psi-\left[\boldsymbol{\nabla}\left(\log\mu\right)\cdot\left(\boldsymbol{\nabla}\psi\right)\right]\right]\sqrt{\mathbf{E}_0\cdot\mathbf{E}_0^*}\hat{\mathbf{e}}_0+\sqrt{\mathbf{E}_0\cdot\mathbf{E}_0^*}\left(\boldsymbol{\nabla}\psi\cdot\boldsymbol{\nabla}\right)\hat{\mathbf{e}}_0$$
$$+\frac{\hat{\mathbf{e}}_0}{2\sqrt{\mathbf{E}_0\cdot\mathbf{E}_0^*}}\left(\boldsymbol{\nabla}\psi\cdot\boldsymbol{\nabla}\right)\left(\mathbf{E}_0\cdot\mathbf{E}_0^*\right)+\left(\sqrt{\mathbf{E}_0\cdot\mathbf{E}_0^*}\hat{\mathbf{e}}_0\cdot\boldsymbol{\nabla}\left(\log n\right)\right)\boldsymbol{\nabla}\psi=0 , \tag{3.132}$$

and thus [see Eq. (3.131)]

$$\left(\boldsymbol{\nabla}\psi\cdot\boldsymbol{\nabla}\right)\hat{\mathbf{e}}_0+\left(\hat{\mathbf{e}}_0\cdot\boldsymbol{\nabla}\left(\log n\right)\right)\boldsymbol{\nabla}\psi=0 , \tag{3.133}$$

or

$$\frac{\mathrm{d}}{\mathrm{d}s}\hat{\mathbf{e}}_0=-\left[\hat{\mathbf{e}}_0\cdot\boldsymbol{\nabla}\left(\log n\right)\right]\hat{\mathbf{s}} , \tag{3.134}$$

thus Eq. (3.129) holds [see Eq. (3.56)]. Note that with the help of Eq. (1.96), which is given by

$$\mathbf{A} \times (\mathbf{B} \times \mathbf{C}) = (\mathbf{A} \cdot \mathbf{C})\,\mathbf{B} - (\mathbf{A} \cdot \mathbf{B})\,\mathbf{C} \;, \tag{3.135}$$

Eq. (3.129) can be rewritten as [as can be seen from Eq. (3.36), $\hat{\mathbf{e}}_0 \cdot \hat{\mathbf{s}} = \mathbf{0}$]

$$\frac{\mathrm{d}}{\mathrm{d}s}\hat{\mathbf{e}}_0 = -\kappa \hat{\mathbf{e}}_0 \times \hat{\mathbf{b}} \;, \tag{3.136}$$

where $\hat{\mathbf{b}} = \hat{\mathbf{s}} \times \hat{\boldsymbol{\nu}}$ is the binormal unit vector.

Using the notation

$$\hat{\mathbf{e}}_0 = e_\nu \hat{\boldsymbol{\nu}} + e_b \hat{\mathbf{b}} \;, \tag{3.137}$$

one finds using Eq. (3.44) that

$$\frac{\mathrm{d}e_\nu}{\mathrm{d}s}\hat{\boldsymbol{\nu}} + \frac{\mathrm{d}e_b}{\mathrm{d}s}\hat{\mathbf{b}} + e_\nu \left(-\kappa \hat{\mathbf{s}} + \tau \hat{\mathbf{b}}\right) - e_b \tau \hat{\boldsymbol{\nu}} = -\kappa e_\nu \hat{\mathbf{s}} \;, \tag{3.138}$$

thus

$$\frac{\mathrm{d}}{\mathrm{d}s}\begin{pmatrix} e_\nu \\ e_b \end{pmatrix} = i\mathcal{K}_g \begin{pmatrix} e_\nu \\ e_b \end{pmatrix} \;, \tag{3.139}$$

where

$$\mathcal{K}_g = \tau \sigma_2 = \tau \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \;. \tag{3.140}$$

The fact that σ_2 is Hermitian, namely $\sigma_2^\dagger = \sigma_2$, ensures that the s evolution of $\hat{\mathbf{e}}_0$ is unitary. The solution of Eq. (3.139) reads

$$\begin{pmatrix} e_\nu(s) \\ e_b(s) \end{pmatrix} = \exp(i\sigma_2\theta) \begin{pmatrix} e_\nu(0) \\ e_b(0) \end{pmatrix} \;, \tag{3.141}$$

where

$$\theta = \int_0^s \mathrm{d}s' \tau(s') \;. \tag{3.142}$$

Using the fact that $\sigma_2^2 = 1$, where 1 denotes the 2×2 identity matrix, one finds that

$$\exp(i\sigma_2\theta) = \sum_{n=0}^{\infty} \frac{(i\sigma_2\theta)^n}{n!} = \cos\theta + i\sigma_2 \sin\theta \;, \tag{3.143}$$

thus

$$\begin{pmatrix} e_\nu(s) \\ e_b(s) \end{pmatrix} = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} e_\nu(0) \\ e_b(0) \end{pmatrix} \;. \tag{3.144}$$

As is shown by Eq. (3.153) below, for the case of a closed curve, i.e. when $\hat{\mathbf{s}}(s) = \hat{\mathbf{s}}(0)$, the following holds

$$\theta = \Omega \;, \tag{3.145}$$

where Ω is the solid angle subtends by the closed curve $\hat{\mathbf{s}}(s')$ at the origin.

Exercise 3.4.2. Calculate the integrated torsion (3.142) for the case of a closed curve, i.e. when $\hat{\mathbf{s}}(s) = \hat{\mathbf{s}}(0)$.

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. $|\mathrm{d}\mathbf{r}/\mathrm{d}s| = 1$. Consider the integrated torsion τ over the following infinitesimal closed curve

$$\mathrm{d}\theta = \left(\int_{\left(s-\frac{\mathrm{d}s}{2},u-\frac{\mathrm{d}u}{2}\right)}^{\left(s+\frac{\mathrm{d}s}{2},u-\frac{\mathrm{d}u}{2}\right)} + \int_{\left(s+\frac{\mathrm{d}s}{2},u-\frac{\mathrm{d}u}{2}\right)}^{\left(s+\frac{\mathrm{d}s}{2},u+\frac{\mathrm{d}u}{2}\right)} + \int_{\left(s+\frac{\mathrm{d}s}{2},u+\frac{\mathrm{d}u}{2}\right)}^{\left(s-\frac{\mathrm{d}s}{2},u+\frac{\mathrm{d}u}{2}\right)} + \int_{\left(s-\frac{\mathrm{d}s}{2},u+\frac{\mathrm{d}u}{2}\right)}^{\left(s-\frac{\mathrm{d}s}{2},u-\frac{\mathrm{d}u}{2}\right)} \right) \tau \mathrm{d}s \mathrm{d}u \; . \tag{3.146}$$

Using the relation

$$\tau = -\hat{\boldsymbol{\nu}} \cdot \frac{\mathrm{d}\hat{\mathbf{b}}}{\mathrm{d}s} = \hat{\mathbf{b}} \cdot \frac{\mathrm{d}\hat{\boldsymbol{\nu}}}{\mathrm{d}s} \; , \tag{3.147}$$

one finds to lowest order in $\mathrm{d}s$ and $\mathrm{d}u$ that

$$\begin{aligned}
\mathrm{d}\theta &= \left[\hat{\mathbf{b}}\left(s,u-\frac{\mathrm{d}u}{2}\right) \cdot \frac{\mathrm{d}\hat{\boldsymbol{\nu}}\left(s,u-\frac{\mathrm{d}u}{2}\right)}{\mathrm{d}s} - \hat{\mathbf{b}}\left(s,u+\frac{\mathrm{d}u}{2}\right) \cdot \frac{\mathrm{d}\hat{\boldsymbol{\nu}}\left(s,u+\frac{\mathrm{d}u}{2}\right)}{\mathrm{d}s} \right] \mathrm{d}s \\
&\quad + \left[\hat{\mathbf{b}}\left(s+\frac{\mathrm{d}s}{2},u\right) \cdot \frac{\mathrm{d}\hat{\boldsymbol{\nu}}\left(s+\frac{\mathrm{d}s}{2},u\right)}{\mathrm{d}u} - \hat{\mathbf{b}}\left(s-\frac{\mathrm{d}s}{2},u\right) \cdot \frac{\mathrm{d}\hat{\boldsymbol{\nu}}\left(s-\frac{\mathrm{d}s}{2},u\right)}{\mathrm{d}u} \right] \mathrm{d}u \\
&= \hat{\mathbf{b}}(s,u) \cdot \left[\frac{\mathrm{d}\hat{\boldsymbol{\nu}}\left(s,u-\frac{\mathrm{d}u}{2}\right)}{\mathrm{d}s} - \frac{\mathrm{d}\hat{\boldsymbol{\nu}}\left(s,u+\frac{\mathrm{d}u}{2}\right)}{\mathrm{d}s} \right] \mathrm{d}s \\
&\quad + \frac{\mathrm{d}\hat{\mathbf{b}}(s,u)}{\mathrm{d}u} \cdot \left[-\frac{\mathrm{d}\hat{\boldsymbol{\nu}}\left(s,u-\frac{\mathrm{d}u}{2}\right)}{\mathrm{d}s} - \frac{\mathrm{d}\hat{\boldsymbol{\nu}}\left(s,u+\frac{\mathrm{d}u}{2}\right)}{\mathrm{d}s} \right] \frac{\mathrm{d}s\mathrm{d}u}{2} \\
&\quad + \hat{\mathbf{b}}(s,u) \cdot \left[\frac{\mathrm{d}\hat{\boldsymbol{\nu}}\left(s+\frac{\mathrm{d}s}{2},u\right)}{\mathrm{d}u} - \frac{\mathrm{d}\hat{\boldsymbol{\nu}}\left(s-\frac{\mathrm{d}s}{2},u\right)}{\mathrm{d}u} \right] \mathrm{d}u \\
&\quad + \frac{\mathrm{d}\hat{\mathbf{b}}(s,u)}{\mathrm{d}s} \cdot \left[\frac{\mathrm{d}\hat{\boldsymbol{\nu}}\left(s+\frac{\mathrm{d}s}{2},u\right)}{\mathrm{d}u} + \frac{\mathrm{d}\hat{\boldsymbol{\nu}}\left(s-\frac{\mathrm{d}s}{2},u\right)}{\mathrm{d}u} \right] \frac{\mathrm{d}s\mathrm{d}u}{2} \\
&= \hat{\mathbf{b}}(s,u) \cdot \left[-\frac{\mathrm{d}^2\hat{\boldsymbol{\nu}}(s,u)}{\mathrm{d}s\mathrm{d}u} + \frac{\mathrm{d}^2\hat{\boldsymbol{\nu}}(s,u)}{\mathrm{d}s\mathrm{d}u} \right] \mathrm{d}s\mathrm{d}u \\
&\quad + \left[\frac{\mathrm{d}\hat{\mathbf{b}}(s,u)}{\mathrm{d}s} \cdot \frac{\mathrm{d}\hat{\boldsymbol{\nu}}(s,u)}{\mathrm{d}u} - \frac{\mathrm{d}\hat{\mathbf{b}}(s,u)}{\mathrm{d}u} \frac{\mathrm{d}\hat{\boldsymbol{\nu}}(s,u)}{\mathrm{d}s} \right] \mathrm{d}s\mathrm{d}u \\
&= \left[\frac{\mathrm{d}\hat{\mathbf{b}}(s,u)}{\mathrm{d}s} \cdot \frac{\mathrm{d}\hat{\boldsymbol{\nu}}(s,u)}{\mathrm{d}u} - \frac{\mathrm{d}\hat{\mathbf{b}}(s,u)}{\mathrm{d}u} \frac{\mathrm{d}\hat{\boldsymbol{\nu}}(s,u)}{\mathrm{d}s} \right] \mathrm{d}s\mathrm{d}u \; .
\end{aligned} \tag{3.148}$$

In general, for a unit vector $\hat{\mathbf{u}}(\alpha)$, where α is a parameter, the following holds

$$\hat{\mathbf{u}} \cdot \frac{\mathrm{d}\hat{\mathbf{u}}}{\mathrm{d}\alpha} = 0 , \tag{3.149}$$

thus only the components in the $\hat{\mathbf{s}}$ direction contribute

$$\mathrm{d}\theta = \left[\left(\hat{\mathbf{s}} \cdot \frac{\mathrm{d}\hat{\mathbf{b}}}{\mathrm{d}s} \right) \left(\hat{\mathbf{s}} \cdot \frac{\mathrm{d}\hat{\boldsymbol{\nu}}}{\mathrm{d}u} \right) - \left(\hat{\mathbf{s}} \cdot \frac{\mathrm{d}\hat{\mathbf{b}}}{\mathrm{d}u} \right) \left(\hat{\mathbf{s}} \cdot \frac{\mathrm{d}\hat{\boldsymbol{\nu}}}{\mathrm{d}s} \right) \right] \mathrm{d}s\mathrm{d}u . \tag{3.150}$$

Moreover, since $\hat{\mathbf{s}} \cdot \hat{\mathbf{b}} = \hat{\mathbf{s}} \cdot \hat{\boldsymbol{\nu}} = 0$, one obtains

$$\begin{aligned} \mathrm{d}\theta &= \left[\left(\frac{\mathrm{d}\hat{\mathbf{s}}}{\mathrm{d}s} \cdot \hat{\mathbf{b}} \right) \left(\frac{\mathrm{d}\hat{\mathbf{s}}}{\mathrm{d}u} \cdot \hat{\boldsymbol{\nu}} \right) - \left(\frac{\mathrm{d}\hat{\mathbf{s}}}{\mathrm{d}u} \cdot \hat{\mathbf{b}} \right) \left(\frac{\mathrm{d}\hat{\mathbf{s}}}{\mathrm{d}s} \cdot \hat{\boldsymbol{\nu}} \right) \right] \mathrm{d}s\mathrm{d}u \\ &= \underbrace{\hat{\mathbf{s}} \cdot \left(\frac{\mathrm{d}\hat{\mathbf{s}}}{\mathrm{d}u} \times \frac{\mathrm{d}\hat{\mathbf{s}}}{\mathrm{d}s} \right)}_{\mathrm{d}A} \mathrm{d}s\mathrm{d}u , \end{aligned} \tag{3.151}$$

where $\mathrm{d}A$ is the area of the infinitesimal closed loop on the unit sphere. This result can be used for evaluating the integral

$$\varDelta\theta = \oint_C \tau \mathrm{d}s\mathrm{d}u \tag{3.152}$$

over a closed loop C by dividing the area into infinitesimally small loops. Thus one concludes that

$$\varDelta\theta = \varOmega , \tag{3.153}$$

where $\varOmega$ is the solid angle that C subtends at the origin.

3.5 Energy Conservation

From Eq. (3.36) one finds that $\mathbf{E}_0$ has no component in the $\hat{\mathbf{s}}$ direction, thus one can write

$$\mathbf{E}_0 = \alpha_{\nu 0} \hat{\boldsymbol{\nu}} + \alpha_{b0} \hat{\mathbf{b}} . \tag{3.154}$$

Using Eq. (3.35) one finds that

$$\mathbf{H}_0 = \sqrt{\frac{\epsilon}{\mu}} \left(\alpha_{\nu 0} \hat{\mathbf{b}} - \alpha_{b0} \hat{\boldsymbol{\nu}} \right) . \tag{3.155}$$

The vectors $\mathbf{E}_0$, $\mathbf{H}_0$, and $\hat{\mathbf{s}}$ are orthogonal to each other. The Poynting phasor vector [See Eq. (2.93)]

$$\begin{aligned}\mathbf{S} &= \frac{c}{8\pi}\mathbf{E}_0 \times \mathbf{H}_0^* \\ &= \frac{c}{8\pi}\sqrt{\frac{\epsilon}{\mu}}\left(\alpha_{\nu 0}\hat{\boldsymbol{\nu}} + \alpha_{b0}\hat{\mathbf{b}}\right) \times \left(\alpha_{\nu 0}^*\hat{\mathbf{b}} - \alpha_{b0}^*\hat{\boldsymbol{\nu}}\right) \\ &= \frac{c}{8\pi}\sqrt{\frac{\epsilon}{\mu}}\left(\left|\alpha_{\nu 0}\right|^2 + \left|\alpha_{b0}\right|^2\right)\hat{\mathbf{s}}\ ,\end{aligned} \tag{3.156}$$

is parallel to $\hat{\mathbf{s}}$ and real and the following holds [See Eq. (2.137)]

$$\epsilon\left|\mathbf{E}_0\right|^2 = \mu\left|\mathbf{H}_0\right|^2 = n\left(8\pi/c\right)\mathbf{S}\cdot\hat{\mathbf{s}} = n\left(8\pi/c\right)\left|\mathbf{S}\right|\ . \tag{3.157}$$

The local time averaged electric and magnetic energy densities are given by [see Eq. (2.137)]

$$\langle w_e\rangle = \frac{\epsilon}{16\pi}\left|\mathbf{E}_0\right|^2\ , \tag{3.158}$$

and

$$\langle w_m\rangle = \frac{\mu}{16\pi}\left|\mathbf{H}_0\right|^2\ , \tag{3.159}$$

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

$$\mathbf{S} = v\left\langle w\right\rangle\hat{\mathbf{s}}\ , \tag{3.160}$$

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 $\mathbf{S}$ and the rate of change in the density of electromagnetic energy u [see Eq. (2.97)].

Claim. In geometrical optics the phasor $\mathbf{S}$ satisfies

$$\boldsymbol{\nabla}\cdot\mathbf{S} = 0\ . \tag{3.161}$$

Proof. Using Eq. (2.149) one finds that

$$\mu\boldsymbol{\nabla}\cdot\left(\frac{1}{\mu}\boldsymbol{\nabla}\psi\right) = \boldsymbol{\nabla}^2\psi - \left[\boldsymbol{\nabla}\left(\log\mu\right)\cdot\left(\boldsymbol{\nabla}\psi\right)\right]\ , \tag{3.162}$$

thus [see Eq. (3.131)]

$$\boldsymbol{\nabla}\cdot\left(\frac{\boldsymbol{\nabla}\psi}{\mu}\right)\left(\mathbf{E}_0\cdot\mathbf{E}_0^*\right) + \left(\frac{\boldsymbol{\nabla}\psi}{\mu}\cdot\boldsymbol{\nabla}\right)\left(\mathbf{E}_0\cdot\mathbf{E}_0^*\right) = 0\ . \tag{3.163}$$

The relations $\boldsymbol{\nabla}\psi = n\hat{\mathbf{s}}$ and $n = \sqrt{\epsilon\mu}$ lead to

$$\boldsymbol{\nabla}\cdot\left(\sqrt{\frac{\epsilon}{\mu}}\hat{\mathbf{s}}\right)(\mathbf{E}_0\cdot\mathbf{E}_0^*)+\left(\sqrt{\frac{\epsilon}{\mu}}\hat{\mathbf{s}}\cdot\boldsymbol{\nabla}\right)(\mathbf{E}_0\cdot\mathbf{E}_0^*)=0\;, \tag{3.164}$$

or

$$\left[(\mathbf{E}_0\cdot\mathbf{E}_0^*)\sqrt{\frac{\epsilon}{\mu}}\right]\boldsymbol{\nabla}\cdot\hat{\mathbf{s}}+\hat{\mathbf{s}}\cdot\boldsymbol{\nabla}\left[(\mathbf{E}_0\cdot\mathbf{E}_0^*)\sqrt{\frac{\epsilon}{\mu}}\right]=0\;, \tag{3.165}$$

thus [see Eq. 2.149]

$$\boldsymbol{\nabla}\cdot\left[(\mathbf{E}_0\cdot\mathbf{E}_0^*)\sqrt{\frac{\epsilon}{\mu}}\hat{\mathbf{s}}\right]=0\;, \tag{3.166}$$

in agreement with Eq. (3.161) [see Eq. (3.157)].

3.5.1 Intensity Along an Optical Ray

The intensity I along an optical ray is defined by

$$I=v\left\langle w\right\rangle\;. \tag{3.167}$$

Exercise 3.5.1. Show that

$$\frac{\mathrm{d}}{\mathrm{d}s}\log I=-\boldsymbol{\nabla}\cdot\hat{\mathbf{s}}\;. \tag{3.168}$$

Solution 3.5.1. With the help of Eqs. (3.160), (3.161) and (3.167) one obtains

$$0=\boldsymbol{\nabla}\cdot\mathbf{S}=\boldsymbol{\nabla}\cdot(I\hat{\mathbf{s}})\;, \tag{3.169}$$

thus [see Eq. (2.149)]

$$0=I\boldsymbol{\nabla}\cdot\hat{\mathbf{s}}+\hat{\mathbf{s}}\cdot\boldsymbol{\nabla}I\;, \tag{3.170}$$

in agreement with Eq. (3.168). Note that with the help of Eq. (3.42) the above result (3.168) can alternatively be written as

$$\frac{\mathrm{d}}{\mathrm{d}s}\left(\log\frac{I}{n}\right)=-\frac{1}{n}\boldsymbol{\nabla}^2\psi\;. \tag{3.171}$$

As can be seen from Eq. (3.168), the evolution of the intensity I along an optical ray depends on the quantity $\boldsymbol{\nabla}\cdot\hat{\mathbf{s}}$. The geometrical meaning of the term $\boldsymbol{\nabla}\cdot\hat{\mathbf{s}}$ is discussed below.

Consider a point P on an optical ray. Let S be the surface of constant ψ (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, $\hat{\mathbf{z}}$ is normal to M and $\hat{\mathbf{x}}'$ and $\hat{\mathbf{y}}'$ are two orthogonal vectors in M. The surface S can be described by a graph

$$(x', y', f(x', y')) \ , \tag{3.172}$$

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

$$\begin{aligned} f\left(\bar{\rho} + \bar{\delta}\right) &= \exp\left(\bar{\delta} \cdot \bar{\nabla}\right) f(\bar{\rho}) \\ &= f(\bar{\rho}) + \left(\bar{\delta} \cdot \bar{\nabla}\right) f(\bar{\rho}) + \frac{1}{2!}\left(\bar{\delta} \cdot \bar{\nabla}\right)^2 f(\bar{\rho}) + \frac{1}{3!}\left(\bar{\delta} \cdot \bar{\nabla}\right)^3 f(\bar{\rho}) + \cdots \ . \end{aligned} \tag{3.173}$$

Thus, to lowest nonvanishing order near $\bar{\rho} = 0$ one has

$$\begin{aligned} f(x', y') &= \frac{1}{2}\left(x' \frac{\partial}{\partial x'} + y' \frac{\partial}{\partial y'}\right)^2 f \\ &= \frac{1}{2}(x', y') \underbrace{\begin{pmatrix} f_{x'x'} & f_{x'y'} \\ f_{y'x'} & f_{y'y'} \end{pmatrix}}_{F} \begin{pmatrix} x' \\ y' \end{pmatrix} \ . \end{aligned} \tag{3.174}$$

The matrix F is symmetric ($F^{\mathrm{T}} = F$), 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 $\hat{\mathbf{x}}$ and $\hat{\mathbf{y}}$ lie in M, $\hat{\mathbf{x}}$ and $\hat{\mathbf{y}}$ are orthogonal to each other and both are orthogonal to $\hat{\mathbf{z}}$. The directions $\hat{\mathbf{x}}$ and $\hat{\mathbf{y}}$ are the principle directions of point P on S. With these coordinates to lowest nonvanishing order the following holds

$$f(x, y) = \frac{1}{2}\left(\frac{x^2}{R_1} + \frac{y^2}{R_2}\right) \ , \tag{3.175}$$

where the principle curvatures are given by

$$R_1 = \frac{1}{f_{xx}} \ , \tag{3.176}$$

and

$$R_2 = \frac{1}{f_{yy}} \ . \tag{3.177}$$

Using this coordinate system the eikonal function can be expressed as

$$\begin{aligned} \psi(\mathbf{r}) &= \exp(\mathbf{r} \cdot \boldsymbol{\nabla}) \psi(0) \\ &= \psi(0) + (\mathbf{r} \cdot \boldsymbol{\nabla}) \psi(0) + \frac{1}{2!}(\mathbf{r} \cdot \boldsymbol{\nabla})^2 \psi(0) + ... \\ &= \psi(0) + x\psi_x + y\psi_y + z\psi_z \\ &\quad + \frac{1}{2}(x, y, z) \begin{pmatrix} \psi_{xx} & \psi_{xy} & \psi_{xz} \\ \psi_{yx} & \psi_{yy} & \psi_{yz} \\ \psi_{zx} & \psi_{zy} & \psi_{zz} \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} + \cdots \ , \end{aligned} \tag{3.178}$$

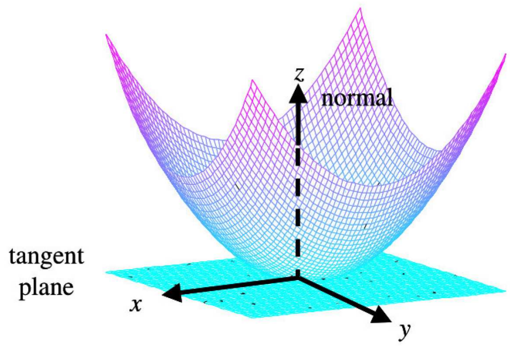

Fig. 3.2. The tangent plane.

On the surface S the eikonal function is constant $\psi = \psi(0)$, thus

$$\psi(x, y, f(x, y)) = \psi(0) \ . \tag{3.179}$$

Substituting this condition in the expansion (3.178) yields

$$\begin{aligned}
&x\psi_x + y\psi_y + \frac{1}{2}\left(\frac{x^2}{R_1} + \frac{y^2}{R_2}\right)\psi_z \\
&+\frac{1}{2}\left(x, y, \frac{1}{2}\left(\frac{x^2}{R_1} + \frac{y^2}{R_2}\right)\right)\begin{pmatrix} \psi_{xx} & \psi_{xy} & \psi_{xz} \\ \psi_{yx} & \psi_{yy} & \psi_{yz} \\ \psi_{zx} & \psi_{zy} & \psi_{zz} \end{pmatrix}\begin{pmatrix} x \\ y \\ \frac{1}{2}\left(\frac{x^2}{R_1} + \frac{y^2}{R_2}\right) \end{pmatrix} + \cdots = 0 \ ,
\end{aligned}$$

thus

$$\psi_x = \psi_y = 0 \ , \tag{3.180}$$

and

$$\psi_z = -R_1\psi_{xx} = -R_2\psi_{yy} \ . \tag{3.181}$$

By using Eq. (2.149) one finds that

$$\begin{aligned}
\boldsymbol{\nabla}\cdot\hat{\mathbf{s}} &= \boldsymbol{\nabla}\cdot\frac{\boldsymbol{\nabla}\psi}{\sqrt{\boldsymbol{\nabla}\psi\cdot\boldsymbol{\nabla}\psi}} = \frac{1}{\sqrt{\boldsymbol{\nabla}\psi\cdot\boldsymbol{\nabla}\psi}}\left(\boldsymbol{\nabla}^2\psi - \frac{1}{2}\frac{\boldsymbol{\nabla}(\boldsymbol{\nabla}\psi\cdot\boldsymbol{\nabla}\psi)}{\boldsymbol{\nabla}\psi\cdot\boldsymbol{\nabla}\psi}\cdot\boldsymbol{\nabla}\psi\right) \\
&= \frac{1}{|\psi_z|}\left(\psi_{xx} + \psi_{yy} + \psi_{zz} - \psi_{zz}\right) \ ,
\end{aligned} \tag{3.182}$$

thus

$$\boldsymbol{\nabla}\cdot\hat{\mathbf{s}} = -\frac{\psi_z}{|\psi_z|}\left(\frac{1}{R_1} + \frac{1}{R_2}\right) \ . \tag{3.183}$$

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. (3.183) the evolution equation for the intensity $I(s)$ along an optical ray (3.171) 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 $\boldsymbol{\nabla}\psi = n\hat{\mathbf{s}}$)

$$\frac{\mathrm{d}\log I}{\mathrm{d}s} = -\frac{1}{R_1(s)} - \frac{1}{R_2(s)} \ . \tag{3.184}$$

As can be seen from the following identity

$$\frac{\frac{\mathrm{d}}{\mathrm{d}s}\left(\frac{R_1R_2}{(R_1+s)(R_2+s)}\right)}{\frac{R_1R_2}{(R_1+s)(R_2+s)}} = -\frac{1}{R_1+s} - \frac{1}{R_2+s} \ , \tag{3.185}$$

the solution can be taken to be given by

$$I(s) = \frac{R_1R_2}{(R_1+s)(R_2+s)} \ , \tag{3.186}$$

where R_1 and R_2 are the principle curvatures for $s = 0$.

The above result (3.186) for the intensity $I(s)$ along an optical ray for the case $n = 1$ can be used in order to express the electric filed $\mathbf{E}(s)$ in geometrical optics along an optical ray [see Eq. (3.20) and note that the relation $\boldsymbol{\nabla}\psi = n\hat{\mathbf{s}}$ for the case of a constant n implies that along the optical ray ψ can be taken to be given by $\psi = ns$]

$$\mathbf{E}(s) = \mathbf{E}(0)\, e^{ik_0 s}\sqrt{\frac{R_1R_2}{(R_1+s)(R_2+s)}} \ . \tag{3.187}$$

3.6 Problems

1. **eikonal approximation in 4-vector formalism** - When the potential 4-vector $A = (\phi, A_1, A_2, A_3)^{\mathrm{T}} = (\phi, \mathbf{A})^{\mathrm{T}}$ [see Eq. (2.26)] is expressed as

   $$A = \mathcal{A}e^{i\Psi} \ , \tag{3.188}$$

   the Maxwell's equation (2.126) becomes

   $$\partial g\left(\partial^{\mathrm{T}}\mathcal{A}^{\mathrm{T}}e^{i\Psi}\eta - \left(\partial^{\mathrm{T}}\mathcal{A}^{\mathrm{T}}e^{i\Psi}\eta\right)^{\mathrm{T}}\right)g = \frac{4\pi}{c}J_{\mathrm{ext}}^{\mathrm{T}} \ . \tag{3.189}$$

   The left hand side of the equation above (3.189) contains a variety of derivative terms. In the eikonal approximation it is assumed that the terms containing derivatives of Ψ are much larger than terms containing derivatives of all other variables (i.e. the metric g (2.112), which depends on the relative permittivity ϵ, on the relative permeability μ, and on the velocity 4-vector U, and the envelope 4-vector $\mathcal{A}$).

a) Show that in the eikonal approximation Eq. (3.189) becomes

$$i\partial g\left(K^{\mathrm{T}}A^{\mathrm{T}}\eta-\eta AK\right)g=\frac{4\pi}{c}J_{\mathrm{ext}}^{\mathrm{T}}\ , \tag{3.190}$$

where

$$K=\partial\Psi=\left(\frac{1}{c}\frac{\partial\Psi}{\partial t},\frac{\partial\Psi}{\partial x_1},\frac{\partial\Psi}{\partial x_2},\frac{\partial\Psi}{\partial x_3}\right)\ . \tag{3.191}$$

b) Show that when the so-called generalized Lorenz gauge condition, which reads

$$\partial g\eta A=0\ , \tag{3.192}$$

is imposed Eq. (3.190) becomes

$$-KgK^{\mathrm{T}}A^{\mathrm{T}}\eta g=\frac{4\pi}{c}J_{\mathrm{ext}}^{\mathrm{T}}\ . \tag{3.193}$$

c) Employ the eikonal approximation to show that when the velocity 3-vector $\mathbf{u}$ vanishes the generalized Lorenz gauge condition (3.192) can be expressed as

$$0=\frac{n^2}{c}\frac{\partial\phi}{\partial t}+\boldsymbol{\nabla}\cdot\mathbf{A}=0\ . \tag{3.194}$$

d) Employ the eikonal approximation to show that when $\mathbf{u}=0$, $J_{\mathrm{ext}}^{\mathrm{T}}=0$ and the generalized Lorenz gauge condition is imposed the Maxwell's equation can be expressed as

$$0=\left(\frac{n^2}{c^2}\frac{\partial^2}{\partial t^2}-\boldsymbol{\nabla}^2\right)A^{\mathrm{T}}\ . \tag{3.195}$$

e) Show that when $J_{\mathrm{ext}}^{\mathrm{T}}=0$ and the generalized Lorenz gauge condition is imposed the Maxwell's equation (3.193) can be expressed as

$$0=K\eta K^{\mathrm{T}}+\frac{n^2-1}{c^2}\left(KU\right)^2\ . \tag{3.196}$$

f) **harmonic time dependency** - Consider the case where the phase factor Ψ has the form

$$\Psi=-\omega t+k_0\psi\left(\mathbf{r}\right)\ , \tag{3.197}$$

where the angular frequency ω is a constant, ψ is time independent, and k_0 is given by

$$k_0=\frac{\omega}{c}\ . \tag{3.198}$$

For this case Eq. (3.191) becomes

$$K = k_0 \left(-1, \boldsymbol{\nabla}\psi\right) \ . \tag{3.199}$$

The real function $\psi(\mathbf{r})$ is commonly called the eikonal function . Express the relation (3.196) for the case where Ψ is given by Eq. (3.197) and the generalized Lorenz gauge condition is imposed.

g) Show that when Ψ is given by Eq. (3.197), the velocity 3-vector $\mathbf{u}$ vanishes and the generalized Lorenz gauge condition is imposed the following holds

$$\mathbf{A} \cdot \boldsymbol{\nabla}\psi = n^2 \phi \ . \tag{3.200}$$

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

$$\mathbf{E} = ik_0 \mathbf{A}_\perp \ , \tag{3.201}$$

$$\mathbf{B} = \boldsymbol{\nabla}\psi \times \mathbf{E} \ , \tag{3.202}$$

where $\mathbf{A}_\perp$ is the component of $\mathbf{A}$ perpendicular to $\boldsymbol{\nabla}\psi$.

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.

$$u_n \boldsymbol{\nabla}^2 \psi + 2 \left(\boldsymbol{\nabla}\psi\right) \cdot \left(\boldsymbol{\nabla} u_n\right) = 0 \ , \tag{3.203}$$

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.

$$\hat{\mathbf{u}} = \frac{\mathbf{u}}{\sqrt{\mathbf{u} \cdot \mathbf{u}^*}} = \left(\hat{u}_1, \hat{u}_2, \hat{u}_3\right) \ . \tag{3.204}$$

a) Show that

$$0 = \left(\boldsymbol{\nabla}\psi\right) \cdot \left(\boldsymbol{\nabla}\hat{u}_n\right) \ . \tag{3.205}$$

b) Contrary to Eq. (3.205), which is obtained from scalar geometrical optics, Eq. (3.129) for the electric field unit vector $\hat{\mathbf{e}}_0$, which is based on the transport equation (3.114), is obtained from a vectorial treatment. Show that these results agree only when the medium is homogeneous.

3. Consider an optical ray striking the interface between two homogeneous materials of refraction indices n_1 and n_2. Show that Snell's law (2.237) (according to which $n_1 \sin\theta_{\mathrm{i}} = n_2 \sin\theta_{\mathrm{t}}$, where θ_{i} and θ_{t} are the angles of incidence and transmission, respectively) can be obtained from the Fermat's principle.
4. Calculate the torsion τ of an helix, which in Cartesian coordinates is given by

$$\mathbf{r}(s) = (r_x, r_y, r_z) = (R\cos ks, R\sin ks, \alpha ks) \ , \tag{3.206}$$

where R, k, and α are real constant and s is a real parameter.

5. Let C be a curve given by a general parametrization $\mathbf{r}(q)$, where $q \in [q_1, q_2]$. Express the curvature κ and the torsion τ in terms of the vectors $\dot{\mathbf{r}} = \mathrm{d}\mathbf{r}/\mathrm{d}q$, $\ddot{\mathbf{r}} = \mathrm{d}^2\mathbf{r}/\mathrm{d}q^2$ and $\dddot{\mathbf{r}} = \mathrm{d}^3\mathbf{r}/\mathrm{d}q^3$.
6. Consider a right circular conic surface whose axis is the z axis, its apex is the origin and its aperture is 2φ. Let $\mathbf{r}(\theta)$, which is given by

$$\mathbf{r}(\theta) = (r_x, r_y, r_z) = r(\theta)\left(\cos\theta, \sin\theta, \frac{1}{\tan\varphi}\right) , \tag{3.207}$$

be a curve on the conic surface, where $\theta \in [\theta_1, \theta_2]$, and $\theta_1 < \theta_2$. Under what condition upon the function $r(\theta)$ the total length from the starting point at $\mathbf{r}_1 = \mathbf{r}(\theta_1)$ to the final one at $\mathbf{r}_2 = \mathbf{r}(\theta_2)$ is locally minimized by the curve $\mathbf{r}(\theta)$?
7. 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 $x_2 = 0$.
   a) Calculate the trajectory of the ray for the case where the refractive index n is given by

   $$n = n_0 \exp(-\gamma x_3) , \tag{3.208}$$

   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 ϕ_0.
   b) Calculate the trajectory of the ray for the case where the refractive index n is given by

   $$n = n_0 \frac{L}{x_3} , \tag{3.209}$$

   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.
8. **Brachistochrone curve** - A particle having mass m moves from point A having Cartesian coordinates $(x, y, z) = (0, 0, 0)$ to point B having Cartesian coordinates $(X, 0, Z)$, where $Z < 0$, under the influence of a uniform gravitational potential given by $U = mgz$. The initial velocity at the starting point A vanishes. The particle moves along a frictionless slide connecting the points A and B. Find a trajectory for the slide that minimizes the travel time from point A to point B.
9. 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 $x_2 = 0$. The ray passes through the origin point $(x_1, x_2, x_3) = (0, 0, 0)$. Express the coordinate x_1 along the ray as a function of the coordinate x_3.

10. **hanging rope** - Consider a rope having a constant mass per unit length λ. The rope is supported at its both ends. Determine the shape of a rope which minimizes its total potential energy in a constant gravitational field having potential given by $V(\mathbf{r}) = gx_3$, where g is the gravitational acceleration on Earth's surface, and x_3 axis is parallel to the gravitational field.
11. 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 $\mathbf{r}(s)$ be an arc-length parameterization of an optical ray traveling in the medium. The variables L^2 and L_3 are defined by

$$L^2 = n^2 r^2 \left[1 - \left(\frac{\mathrm{d}r}{\mathrm{d}s}\right)^2\right] , \tag{3.210}$$

$$L_3 = nr^2 \sin^2\theta \frac{\mathrm{d}\phi}{\mathrm{d}s} , \tag{3.211}$$

where, in general, the spherical coordinates (r, θ, ϕ) are related to the Cartesian coordinates (x, y, z) by

$$(x, y, z) = r(\sin\theta\cos\phi, \sin\theta\sin\phi, \cos\theta) . \tag{3.212}$$

Calculate the derivatives $\mathrm{d}L^2/\mathrm{d}s$ and $\mathrm{d}L_3/\mathrm{d}s$ along the optical ray.
12. 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

$$n(r) = \sqrt{2 - \left(\frac{r}{R}\right)^2} , \tag{3.213}$$

    where R is a positive constant (Luneburg lens). Assume that $r \leq R$.
    b) $n(r)$ is given by

$$n(r) = \frac{n_0}{1 + \left(\frac{r}{R}\right)^2} , \tag{3.214}$$

    where n_0 and R are positive constants (Maxwell's fish-eye).
13. **Gravitational lensing** - According to the general theory of relativity a gravitational field gives rise to deflection of optical rays. Consider an optical ray traveling in vacuum in the presence of a gravitational filed $\Phi(\mathbf{r})$. For the case of a weak gravitational field (i.e. to first order in Φ) the gravity-induced deflection can be evaluated from the ray equation of geometrical optics for rays traveling in a medium having refractive index given by

$$n(\mathbf{r}) = 1 - \frac{2\Phi(\mathbf{r})}{c^2} . \tag{3.215}$$

For the case of a point mass M the gravitational potential Φ is given by

$$\Phi(\mathbf{r}) = -\frac{GM}{r} \ ,$$

where $G = 6.67259 \times 10^{-11}\,\mathrm{m}^3\,\mathrm{kg}^{-1}\,\mathrm{s}^{-2}$ is the gravitational constant and where $r = |\mathbf{r}|$. Calculate the deflection angle α of an optical ray travelling near a point mass. Express the result in terms of the smallest distance r_0 between the optical ray and the point mass.

14. **Reflection off a moving mirror** - Consider a mirror moving with respect to an inertial frame denoted by S' at a constant velocity given by βc, where $0 \leq \beta < 1$ and c is the speed of light.
    a) Find a relation between the angle of incidence θ'_{i} and the angle of reflection θ'_{r}. Assume that in the rest frame of the mirror, which is denoted by S, the so-called law of reflection (2.236), according to which $\theta_{\mathrm{i}} = \theta_{\mathrm{r}}$, holds ($\theta_{\mathrm{i}}$ and θ_{r} are the angles of incidence and reflection, respectively, as being measure in S).
    b) Show that the result obtained in (a), i.e. the relation between θ'_{i} and θ'_{r}, is consistent with the Fermat's principle.
15. Eliminating $\mathbf{H}_{m+1}$ from Eq. (3.23) and substituting into Eq. (3.22) yield

$$-\mu \boldsymbol{\nabla} \times \mathbf{H}_m - \boldsymbol{\nabla}\psi \times (\boldsymbol{\nabla} \times \mathbf{E}_m) = n^2 \mathbf{E}_{m+1} + (\boldsymbol{\nabla}\psi) \times [(\boldsymbol{\nabla}\psi) \times \mathbf{E}_{m+1}] \ , \tag{3.216}$$

thus, for the case $m = 0$ one has

$$-\mu \boldsymbol{\nabla} \times \mathbf{H}_0 - \boldsymbol{\nabla}\psi \times (\boldsymbol{\nabla} \times \mathbf{E}_0) = n^2 \mathbf{E}_1 + (\boldsymbol{\nabla}\psi) \times [(\boldsymbol{\nabla}\psi) \times \mathbf{E}_1] \ . \tag{3.217}$$

Use the above result (3.217) in order to derive the transport equation (3.114).

16. Show that

$$\hat{\boldsymbol{\nu}} \cdot \boldsymbol{\nabla} \times \hat{\mathbf{s}} = 0 \ , \tag{3.218}$$

and

$$\hat{\mathbf{b}} \cdot \boldsymbol{\nabla} \times \hat{\mathbf{s}} = \kappa \ . \tag{3.219}$$

17. Show that the eccentricity e of the polarization ellipse [see Eq. (2.162)] is a constant along an optical ray.
18. Consider a medium whose refractive index n is given in Cartesian coordinates (x, y, z) by

$$n = n_0 \sqrt{1 - K^2 (x^2 + y^2)} \ , \tag{3.220}$$

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

$$x^2(s) + y^2(s) = R^2 , \tag{3.221}$$

where R is a positive constants. Let $\hat{\mathbf{e}}_0(s)$ be a unit vector pointing in the direction of the electric field. The polarization is assumed to be rectilinear, i.e. the unit vector $\hat{\mathbf{e}}_0(s)$ is assumed to be real. Let $\phi_{\mathrm{e}}(s)$ be the angle between $\hat{\mathbf{e}}_0(s)$ and the binormal unit vector $\hat{\mathbf{b}}$. Calculate $\mathrm{d}\phi_{\mathrm{e}}/\mathrm{d}s$.

19. 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

$$S = -mc^2 \int \mathrm{d}\tau - \frac{q}{c} \int \mathrm{d}\tau \, U^{\mathrm{T}} \eta A , \tag{3.222}$$

where τ is the proper time [see Eq. (1.13)], $U = \mathrm{d}X/\mathrm{d}\tau$ is the velocity 4-vector of the particle [see Eq. (1.64)], η is the Minkowski metric (1.14) and $A = (\phi, A_1, A_2, A_3)^{\mathrm{T}} = (\phi, \mathbf{A})^{\mathrm{T}}$ is the electromagnetic potential 4-vector (2.26). Note that the integral along the trajectory $\int \mathrm{d}\tau$ 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 $U^{\mathrm{T}} \eta A$ is Lorentz invariant, since

$$(U')^{\mathrm{T}} \eta A' = U^{\mathrm{T}} \Lambda^{\mathrm{T}} \eta \Lambda A = U^{\mathrm{T}} \eta A . \tag{3.223}$$

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 $\mathbf{r}_1$ at time t_1 to point $\mathbf{r}_2$ 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.

20. 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 $U = \mathrm{d}X/\mathrm{d}\tau$, where τ is the proper time.
    b) Consider the case where in a frame commoving with the particle the electric $\mathbf{E}$ and magnetic $\mathbf{B}$ fields are given by $\mathbf{E} = E_1 \hat{\mathbf{x}}_1$ and $\mathbf{B} = B_1 \hat{\mathbf{x}}_1$, 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 $\partial^{\mathrm{T}} \mathcal{A}^{\mathrm{T}} e^{i\Psi}$ one finds that

$$\partial^{\mathrm{T}} A^{\mathrm{T}} = \partial^{\mathrm{T}} \mathcal{A} e^{i\Psi} \simeq i e^{i\Psi} K^{\mathrm{T}} \mathcal{A}^{\mathrm{T}} = i K^{\mathrm{T}} A^{\mathrm{T}} . \tag{3.224}$$

a) With the help of Eq. (3.224) one finds that Eq. (3.189) becomes Eq. (3.190).
b) When generalized Lorenz gauge condition (3.192) is imposed Eq. (3.190) becomes

$$i\partial g K^{\mathrm{T}} A^{\mathrm{T}} \eta g = \frac{4\pi}{c} J_{\mathrm{ext}}^{\mathrm{T}} \; . \tag{3.225}$$

In the eikonal approximation Eq. (3.225) can be replaced by Eq. (3.193) [see Eqs. (3.188), (3.224) and (3.191), and note that it is assumed that terms containing derivatives of K can be neglected as well]. Moreover, in the eikonal approximation Eq. (3.192) can be replaced by

$$K g \eta A = 0 \; . \tag{3.226}$$

c) With the help of the eikonal approximation and Eq. (2.112), which reads $g = \mu^{-1/2}\left(\eta + \left(\xi/c^2\right) U U^{\mathrm{T}}\right)$, one finds that Eq. (3.192) can be rewritten as

$$\partial A + \frac{\xi}{c^2} \partial U U^{\mathrm{T}} \eta A = 0 \; . \tag{3.227}$$

For a vanishing $\mathbf{u}$ the velocity 4-vector U becomes $U = (c, 0, 0, 0)^{\mathrm{T}}$ [see Eq. (2.113)], and thus Eq. (3.194) holds (recall that $\xi = n^2 - 1$).
d) With the help of the eikonal approximation one finds that for this case [see Eqs. (2.112) and (3.224)]

$$\begin{aligned} i\partial g K^{\mathrm{T}} A^{\mathrm{T}} &= \frac{1}{\sqrt{\mu}} \partial \left(\eta + \frac{\xi}{c^2} U U^{\mathrm{T}}\right) \partial^{\mathrm{T}} A^{\mathrm{T}} \\ &= \frac{1}{\sqrt{\mu}} \left[\partial \eta \partial^{\mathrm{T}} + \left(n^2 - 1\right) \frac{1}{c^2} \frac{\partial^2}{\partial t^2}\right] A^{\mathrm{T}} \\ &= \frac{1}{\sqrt{\mu}} \left(\frac{n^2}{c^2} \frac{\partial^2}{\partial t^2} - \boldsymbol{\nabla}\right) A^{\mathrm{T}} \; , \end{aligned} \tag{3.228}$$

and thus Eq. (3.225) leads to Eq. (3.195).
e) With the help of Eq. (2.112) one finds that

$$K g K^{\mathrm{T}} = \frac{1}{\sqrt{\mu}} \left(K \eta K^{\mathrm{T}} + \frac{n^2 - 1}{c^2} K U U^{\mathrm{T}} K^{\mathrm{T}}\right) \; , \tag{3.229}$$

and thus Eq. (3.193) leads to Eq. (3.196).
f) By expressing the velocity 4-vector as $U = \gamma c (1, \boldsymbol{\beta})^{\mathrm{T}}$, where $\boldsymbol{\beta}$ is related to the velocity 3-vector $\mathbf{u}$ by $\boldsymbol{\beta} = \mathbf{u}/c$ [see Eq. (2.113)], one finds that Eq. (3.196) becomes [see Eq. (3.199)]

$$0 = 1 - (\boldsymbol{\nabla}\psi)^2 + \frac{\left(n^2 - 1\right)(1 - \boldsymbol{\beta} \cdot \boldsymbol{\nabla}\psi)^2}{1 - \beta^2} \; , \tag{3.230}$$

or

$$(\boldsymbol{\nabla}\psi)^2 = n^2 + \frac{\left(n^2-1\right)\left[(1-\boldsymbol{\beta}\cdot\boldsymbol{\nabla}\psi)^2 - \left(1-\beta^2\right)\right]}{1-\beta^2} \; . \tag{3.231}$$

Note that when $\boldsymbol{\beta} = 0$ Eq. (3.231) yields $(\boldsymbol{\nabla}\psi)^2 = n^2$.

g) For this case Eq. (3.226) becomes [see Eqs. (2.125) and (3.199)]

$$K \begin{pmatrix} n^2 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} A = 0 \; , \tag{3.232}$$

thus Eq. (3.200) holds.

h) With the help of Eqs. (2.29) and (3.224) one finds that in the eikonal approximation Eq. (2.28) becomes

$$iK^{\mathrm{T}}A^{\mathrm{T}}\eta - \left(iK^{\mathrm{T}}A^{\mathrm{T}}\eta\right)^{\mathrm{T}} = \begin{pmatrix} 0 & E_1 & E_2 & E_3 \\ -E_1 & 0 & -B_3 & B_2 \\ -E_2 & B_3 & 0 & -B_1 \\ -E_3 & -B_2 & B_1 & 0 \end{pmatrix} \; , \tag{3.233}$$

or [see Eqs. (2.26) and (3.199)]

$$\begin{gathered} ik_0 \begin{pmatrix} 0 & A_1 - \frac{\partial\psi}{\partial x_1}\phi & A_2 - \frac{\partial\psi}{\partial x_2}\phi & A_3 - \frac{\partial\psi}{\partial x_3}\phi \\ \frac{\partial\psi}{\partial x_1}\phi - A_1 & 0 & -\frac{\partial\psi}{\partial x_1}A_2 + \frac{\partial\psi}{\partial x_2}A_1 & -\frac{\partial\psi}{\partial x_1}A_3 + \frac{\partial\psi}{\partial x_3}A_1 \\ \frac{\partial\psi}{\partial x_2}\phi - A_2 & -\frac{\partial\psi}{\partial x_2}A_1 + \frac{\partial\psi}{\partial x_1}A_2 & 0 & -\frac{\partial\psi}{\partial x_2}A_3 + \frac{\partial\psi}{\partial x_3}A_2 \\ \frac{\partial\psi}{\partial x_3}\phi - A_3 & -\frac{\partial\psi}{\partial x_3}A_1 + \frac{\partial\psi}{\partial x_1}A_3 & -\frac{\partial\psi}{\partial x_3}A_2 + \frac{\partial\psi}{\partial x_2}A_3 & 0 \end{pmatrix} \\ = \begin{pmatrix} 0 & E_1 & E_2 & E_3 \\ -E_1 & 0 & -B_3 & B_2 \\ -E_2 & B_3 & 0 & -B_1 \\ -E_3 & -B_2 & B_1 & 0 \end{pmatrix} \; , \end{gathered} \tag{3.234}$$

and thus

$$\mathbf{E} = ik_0\left(\mathbf{A} - \phi\boldsymbol{\nabla}\psi\right) \; , \tag{3.235}$$

and

$$\mathbf{B} = ik_0\left(\boldsymbol{\nabla}\psi \times \mathbf{A}\right) \; . \tag{3.236}$$

With the help of the gauge condition $\mathbf{A}\cdot\boldsymbol{\nabla}\psi = n^2\phi$ (3.200) Eq. (3.235) can be rewritten as

$$\mathbf{E} = ik_0\left(\mathbf{A} - \frac{\mathbf{A}\cdot\boldsymbol{\nabla}\psi}{n^2}\boldsymbol{\nabla}\psi\right) \; . \tag{3.237}$$

According to the eikonal equation (3.7) $(\boldsymbol{\nabla}\psi)^2 = n^2$, and thus $\mathbf{E} = ik_0\mathbf{A}_\perp$, in agreement with Eq. (3.201), where $\mathbf{A}_\perp = \mathbf{A} - \mathbf{A}_\parallel$ is the component of $\mathbf{A}$ perpendicular to $\boldsymbol{\nabla}\psi$ and where $\mathbf{A}_\parallel = n^{-2}(\mathbf{A}\cdot\boldsymbol{\nabla}\psi)\boldsymbol{\nabla}\psi$ is the component of $\mathbf{A}$ parallel to $\boldsymbol{\nabla}\psi$. With the help of the above result (3.201) Eq. (3.236) becomes Eq. (3.202). Thus, both fields $\mathbf{E}$ and $\mathbf{B}$ are perpendicular to $\boldsymbol{\nabla}\psi$, and they are also perpendicular to each other.

2. With the help of Eq. (3.203) and the relation $u_n = \sqrt{\mathbf{u}\cdot\mathbf{u}^*}\hat{u}_n$ [see Eq. (3.204)] one finds that

$$\begin{aligned} 0 &= \sqrt{\mathbf{u}\cdot\mathbf{u}^*}\hat{u}_n\boldsymbol{\nabla}^2\psi + 2(\boldsymbol{\nabla}\psi)\cdot\left(\boldsymbol{\nabla}\left(\sqrt{\mathbf{u}\cdot\mathbf{u}^*}\hat{u}_n\right)\right) \\ &= 2\hat{u}_n\left[\frac{\sqrt{\mathbf{u}\cdot\mathbf{u}^*}}{2}\boldsymbol{\nabla}^2\psi + (\boldsymbol{\nabla}\psi)\cdot\boldsymbol{\nabla}\left(\sqrt{\mathbf{u}\cdot\mathbf{u}^*}\right)\right] + 2\sqrt{\mathbf{u}\cdot\mathbf{u}^*}(\boldsymbol{\nabla}\psi)\cdot(\boldsymbol{\nabla}\hat{u}_n)\ . \end{aligned} \tag{3.238}$$

a) The following holds

$$\begin{aligned} 2\sqrt{\mathbf{u}\cdot\mathbf{u}^*}(\boldsymbol{\nabla}\psi)\cdot\boldsymbol{\nabla}\sqrt{\mathbf{u}\cdot\mathbf{u}^*} &= (\boldsymbol{\nabla}\psi)\cdot\boldsymbol{\nabla}(\mathbf{u}\cdot\mathbf{u}^*) \\ &= \sum_{n=1}^{3}(\boldsymbol{\nabla}\psi)\cdot\boldsymbol{\nabla}(u_n u_n^*) \\ &= \sum_{n=1}^{3}(\boldsymbol{\nabla}\psi)\cdot(u_n^*\boldsymbol{\nabla}u_n + u_n\boldsymbol{\nabla}u_n^*)\ , \end{aligned} \tag{3.239}$$

where [see Eq. (3.203)]

$$\sum_{n=1}^{3}(\boldsymbol{\nabla}\psi)\cdot(u_n^*\boldsymbol{\nabla}u_n + u_n\boldsymbol{\nabla}u_n^*) = -(\mathbf{u}\cdot\mathbf{u}^*)\boldsymbol{\nabla}^2\psi\ , \tag{3.240}$$

and thus

$$(\boldsymbol{\nabla}\psi)\cdot\boldsymbol{\nabla}\sqrt{\mathbf{u}\cdot\mathbf{u}^*} = -\frac{\sqrt{\mathbf{u}\cdot\mathbf{u}^*}}{2}\boldsymbol{\nabla}^2\psi\ . \tag{3.241}$$

The above results (3.238) and (3.241) yield together to Eq. (3.205).

b) With the help of the relation

$$\frac{\mathrm{d}}{\mathrm{d}s} = \frac{\boldsymbol{\nabla}\psi\cdot\boldsymbol{\nabla}}{n}\ ,$$

one finds that Eq. (3.205) can be rewritten as

$$\frac{\mathrm{d}\hat{\mathbf{u}}}{\mathrm{d}s} = 0\ , \tag{3.242}$$

whereas Eq. (3.129) can be written as [see Eq. (3.134)]

$$\frac{\mathrm{d}\hat{\mathbf{e}}_0}{\mathrm{d}s} = -\left[\hat{\mathbf{e}}_0 \cdot \boldsymbol{\nabla}\left(\log n\right)\right]\hat{\mathbf{s}} \,, \tag{3.243}$$

thus only when the medium is homogeneous agreement is obtained.

3. The interface between the materials is taken to be the plane $z = 0$. Consider a transmission process from the initial point $\mathbf{r}_1 = (-\alpha L, 0, L)$ inside the material of refracting index n_1 to the point $\mathbf{r}_2 = (\alpha L, 0, -L)$ inside the material of refractive index n_2, where $\alpha > 0$ is dimensionless and $L > 0$ is the distance between both points $\mathbf{r}_1$ and $\mathbf{r}_2$ and the interface. The optical ray is assumed to be made of two straight sections (explain why), the first from the point $\mathbf{r}_1$ to a point on the interface $\mathbf{r}_0 = (\eta L, 0, 0)$, where η is a dimensionless real number, and the second is from the point $\mathbf{r}_0$ to the point $\mathbf{r}_2$. The total time of flight T can be expressed as

$$cT = n_1 \left|\mathbf{r}_1 - \mathbf{r}_0\right| + n_2 \left|\mathbf{r}_2 - \mathbf{r}_0\right| \ , \tag{3.244}$$

where

$$\frac{\mathbf{r}_1 - \mathbf{r}_0}{\left|\mathbf{r}_1 - \mathbf{r}_0\right|} = \frac{\left(-\left(\alpha+\eta\right)L, 0, L\right)}{L\sqrt{\left(\alpha+\eta\right)^2+1}} = \left(-\sin\theta_{\mathrm{i}}, 0, \cos\theta_{\mathrm{i}}\right) \ , \tag{3.245}$$

$$\frac{\mathbf{r}_2 - \mathbf{r}_0}{\left|\mathbf{r}_2 - \mathbf{r}_0\right|} = \frac{\left(\left(\alpha-\eta\right)L, 0, L\right)}{L\sqrt{\left(\alpha-\eta\right)^2+1}} = \left(\sin\theta_{\mathrm{t}}, 0, \cos\theta_{\mathrm{t}}\right) \ . \tag{3.246}$$

The requirement that T is minimized leads to

$$0 = \frac{\mathrm{d}T}{\mathrm{d}\eta} = \frac{L}{c}\left(\frac{n_1\left(\alpha+\eta\right)}{\sqrt{\left(\alpha+\eta\right)^2+1}} - \frac{n_2\left(\alpha-\eta\right)}{\sqrt{\left(\alpha-\eta\right)^2+1}}\right) \ , \tag{3.247}$$

and thus

$$n_1 \sin\theta_{\mathrm{i}} = n_2 \sin\theta_{\mathrm{t}} \ . \tag{3.248}$$

4. The following holds

$$\frac{\mathrm{d}\mathbf{r}}{\mathrm{d}s} = k\left(-R\sin ks, R\cos ks, \alpha\right) \ . \tag{3.249}$$

The condition for s to be an arc-length parameter reads

$$1 = \left|\frac{\mathrm{d}\mathbf{r}}{\mathrm{d}s}\right| = k\sqrt{R^2+\alpha^2} \,, \tag{3.250}$$

or

$$k = \frac{1}{\sqrt{R^2+\alpha^2}} \ . \tag{3.251}$$

The following holds

$$\hat{\mathbf{s}} = \frac{\mathrm{d}\mathbf{r}}{\mathrm{d}s} = \frac{1}{\sqrt{R^2+\alpha^2}}\left(-R\sin\frac{s}{\sqrt{R^2+\alpha^2}}, R\cos\frac{s}{\sqrt{R^2+\alpha^2}}, \alpha\right) , \tag{3.252}$$

and

$$\frac{\mathrm{d}\hat{\mathbf{s}}}{\mathrm{d}s} = \frac{R}{R^2+\alpha^2}\left(-\cos\frac{s}{\sqrt{R^2+\alpha^2}}, -\sin\frac{s}{\sqrt{R^2+\alpha^2}}, 0\right) , \tag{3.253}$$

thus

$$\kappa = \frac{R}{R^2+\alpha^2} , \tag{3.254}$$

and

$$\hat{\boldsymbol{\nu}} = -\left(\cos\frac{s}{\sqrt{R^2+\alpha^2}}, \sin\frac{s}{\sqrt{R^2+\alpha^2}}, 0\right) . \tag{3.255}$$

Moreover, one has

$$\begin{aligned}\hat{\mathbf{b}} = \hat{\mathbf{s}}\times\hat{\boldsymbol{\nu}} &= \frac{1}{\sqrt{R^2+\alpha^2}}\begin{vmatrix} \mathbf{x} & \mathbf{y} & \mathbf{z} \\ -R\sin\frac{s}{\sqrt{R^2+\alpha^2}} & R\cos\frac{s}{\sqrt{R^2+\alpha^2}} & \alpha \\ -\cos\frac{s}{\sqrt{R^2+\alpha^2}} & -\sin\frac{s}{\sqrt{R^2+\alpha^2}} & 0 \end{vmatrix} \\ &= \frac{1}{\sqrt{R^2+\alpha^2}}\left(\alpha\sin\frac{s}{\sqrt{R^2+\alpha^2}}, -\alpha\cos\frac{s}{\sqrt{R^2+\alpha^2}}, R\right) ,\end{aligned} \tag{3.256}$$

thus

$$\begin{aligned}\frac{\mathrm{d}\hat{\mathbf{b}}}{\mathrm{d}s} &= -\frac{\alpha}{R^2+\alpha^2}\left(-\cos\frac{s}{\sqrt{R^2+\alpha^2}}, -\sin\frac{s}{\sqrt{R^2+\alpha^2}}, 0\right) \\ &= -\tau\hat{\boldsymbol{\nu}} ,\end{aligned} \tag{3.257}$$

where the torsion τ is given by

$$\tau = \frac{\alpha}{R^2+\alpha^2}. \tag{3.258}$$

5. Let $\mathbf{r}(s)$ be an arc-length parametrization of the same curve. The curvature κ is given by [see Eqs. (3.43) and (3.44)]

$$\kappa = \left|\frac{\mathrm{d}\mathbf{r}}{\mathrm{d}s}\times\frac{\mathrm{d}^2\mathbf{r}}{\mathrm{d}s^2}\right| = \left|\frac{\mathrm{d}^2\mathbf{r}}{\mathrm{d}s^2}\right| , \tag{3.259}$$

and the torsion τ is given by [see Eq. (3.50)]

$$\tau = \hat{\boldsymbol{\nu}}\cdot\frac{\mathrm{d}\left(\hat{\boldsymbol{\nu}}\times\hat{\mathbf{s}}\right)}{\mathrm{d}s} , \tag{3.260}$$

where $\hat{\mathbf{s}} = \mathrm{d}\mathbf{r}/\mathrm{d}s$ and $\hat{\boldsymbol{\nu}} = \kappa^{-1}\left(\mathrm{d}^2\mathbf{r}/\mathrm{d}s^2\right)$, and thus [see Eqs. (3.44) and (2.324)]

$$\tau = \hat{\boldsymbol{\nu}} \cdot \frac{\mathrm{d}\hat{\boldsymbol{\nu}}}{\mathrm{d}s} \times \hat{\mathbf{s}} = \frac{\frac{\mathrm{d}\mathbf{r}}{\mathrm{d}s} \cdot \left(\frac{\mathrm{d}^2\mathbf{r}}{\mathrm{d}s^2} \times \frac{\mathrm{d}^3\mathbf{r}}{\mathrm{d}s^3}\right)}{\left|\frac{\mathrm{d}^2\mathbf{r}}{\mathrm{d}s^2}\right|^2} . \tag{3.261}$$

Using the relation

$$\frac{\mathrm{d}}{\mathrm{d}s} = \frac{\mathrm{d}q}{\mathrm{d}s}\frac{\mathrm{d}}{\mathrm{d}q} = \frac{1}{|\dot{\mathbf{r}}|}\frac{\mathrm{d}}{\mathrm{d}q} , \tag{3.262}$$

one finds that in general parametrization Eq. (3.259) becomes [recall that $\mathbf{A} \times \mathbf{A} = 0$ for any vector $\mathbf{A}$]

$$\kappa = \frac{|\dot{\mathbf{r}} \times \ddot{\mathbf{r}}|}{|\dot{\mathbf{r}}|^3} , \tag{3.263}$$

and Eq. (3.261) becomes [see Eq. (2.324)]

$$\tau = \frac{\frac{\mathrm{d}\mathbf{r}}{\mathrm{d}s} \cdot \left(\frac{\mathrm{d}^2\mathbf{r}}{\mathrm{d}s^2} \times \frac{\mathrm{d}^3\mathbf{r}}{\mathrm{d}s^3}\right)}{\left|\frac{\mathrm{d}^2\mathbf{r}}{\mathrm{d}s^2}\right|^2} = \frac{\frac{\mathrm{d}^3\mathbf{r}}{\mathrm{d}s^3} \cdot \left(\frac{\mathrm{d}\mathbf{r}}{\mathrm{d}s} \times \frac{\mathrm{d}^2\mathbf{r}}{\mathrm{d}s^2}\right)}{\left(\frac{|\dot{\mathbf{r}} \times \ddot{\mathbf{r}}|}{|\dot{\mathbf{r}}|^3}\right)^2} . \tag{3.264}$$

The following holds

$$\frac{\mathrm{d}\mathbf{r}}{\mathrm{d}s} \times \frac{\mathrm{d}^2\mathbf{r}}{\mathrm{d}s^2} = \frac{\dot{\mathbf{r}}}{|\dot{\mathbf{r}}|} \times \frac{\mathrm{d}\frac{\dot{\mathbf{r}}}{|\dot{\mathbf{r}}|}}{\mathrm{d}s} = \frac{\dot{\mathbf{r}} \times \ddot{\mathbf{r}}}{|\dot{\mathbf{r}}|^3} , \tag{3.265}$$

and thus [again, recall that $\mathbf{A} \times \mathbf{A} = 0$ for any vector $\mathbf{A}$]

$$\tau = \frac{|\dot{\mathbf{r}}|^3 \frac{\mathrm{d}^3\mathbf{r}}{\mathrm{d}s^3} \cdot (\dot{\mathbf{r}} \times \ddot{\mathbf{r}})}{|\dot{\mathbf{r}} \times \ddot{\mathbf{r}}|^2} = \frac{\dddot{\mathbf{r}} \cdot (\dot{\mathbf{r}} \times \ddot{\mathbf{r}})}{|\dot{\mathbf{r}} \times \ddot{\mathbf{r}}|^2} , \tag{3.266}$$

or [see Eq. (2.324)]

$$\tau = \frac{\dot{\mathbf{r}} \cdot (\ddot{\mathbf{r}} \times \dddot{\mathbf{r}})}{|\dot{\mathbf{r}} \times \ddot{\mathbf{r}}|^2} . \tag{3.267}$$

6. The length of the curve l is given by

$$l = \int_{\theta_1}^{\theta_2} \mathrm{d}\theta \, \mathcal{L} , \tag{3.268}$$

where

$$
\begin{aligned}
\mathcal{L} &= \left| \frac{\mathrm{d}\mathbf{r}}{\mathrm{d}\theta} \right| \\
&= \left| r' \left(\cos\theta, \sin\theta, \frac{1}{\tan\varphi} \right) + r\left(-\sin\theta, \cos\theta, 0\right) \right| \\
&= \sqrt{r^2 + \left(\frac{r'}{\sin\varphi} \right)^2} ,
\end{aligned}
\tag{3.269}
$$

and where $r' = \mathrm{d}r/\mathrm{d}\theta$. The total length is locally minimized provided that [see Eq. (3.85)]

$$
\frac{\partial \mathcal{L}}{\partial r} = \frac{\mathrm{d}}{\mathrm{d}\theta} \frac{\partial \mathcal{L}}{\partial r'} , \tag{3.270}
$$

thus

$$
\frac{r}{\sqrt{r^2 + \left(\frac{r'}{\sin\varphi} \right)^2}} = \frac{\mathrm{d}}{\mathrm{d}\theta} \frac{\frac{r'}{\sin^2\varphi}}{\sqrt{r^2 + \left(\frac{r'}{\sin\varphi} \right)^2}} . \tag{3.271}
$$

In terms of the variable q, which is defined by

$$
\tan q = \frac{r'}{r \sin\varphi} , \tag{3.272}
$$

Eq. (3.271) becomes

$$
\frac{\sin\varphi}{\sqrt{1 + \tan^2 q}} = \frac{\mathrm{d}}{\mathrm{d}\theta} \frac{\tan q}{\sqrt{1 + \tan^2 q}} , \tag{3.273}
$$

and thus by using the identity

$$
\frac{\mathrm{d}}{\mathrm{d}q} \frac{\tan q}{\sqrt{1 + \tan^2 q}} = \frac{1}{\sqrt{1 + \tan^2 q}} , \tag{3.274}
$$

one obtains

$$
\frac{\mathrm{d}q}{\mathrm{d}\theta} = \sin\varphi . \tag{3.275}
$$

The solution is given by

$$
q = (\theta_0 + \theta) \sin\varphi . \tag{3.276}
$$

where θ_0 is a constant, or in terms of r and r' by

$$
\tan\left((\theta_0 + \theta) \sin\varphi \right) = \frac{r'}{r \sin\varphi} . \tag{3.277}
$$

Integration yields

$$r=\frac{r_0}{\cos\left((\theta_0+\theta)\sin\varphi\right)}\;. \tag{3.278}$$

where r_0 is a constant. Note that the above result (3.278) can be alternatively obtained by unfolding the conic surface into a planar circular sector (having angle $\phi = 2\pi\sin\varphi$). Straight lines are the curves that minimize length between two given points on the planar circular sector. Mapping such straight lines back to the conic surface yields Eq. (3.278).

7. Let $\mathbf{x}_{\mathrm{C}}(q) = (x_1(q), x_2(q), x_3(q))$ 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 (3.94) yield

$$0=\frac{\mathrm{d}}{\mathrm{d}x_1}\frac{n}{\sqrt{\dot{\mathbf{x}}_{\mathrm{C}}^2}}\;, \tag{3.279}$$

$$\sqrt{\dot{\mathbf{x}}_{\mathrm{C}}^2}\frac{\partial n}{\partial x_3}=\frac{\mathrm{d}}{\mathrm{d}x_1}\frac{nf}{\sqrt{\dot{\mathbf{x}}_{\mathrm{C}}^2}}\;, \tag{3.280}$$

where

$$\dot{\mathbf{x}}_{\mathrm{C}}^2=1+f^2\;, \tag{3.281}$$

and where

$$f=\frac{\mathrm{d}x_3}{\mathrm{d}x_1}\;, \tag{3.282}$$

and thus the following holds

$$\frac{\partial\log n}{\partial x_3}\left(1+f^2\right)=\frac{\mathrm{d}f}{\mathrm{d}x_1}\;. \tag{3.283}$$

a) For this case $\partial\log n/\partial x_3 = -\gamma$. Integrating Eq. (3.283) yields (recall the initial condition $f = \mathrm{d}x_3/\mathrm{d}x_1 = \tan\phi_0$ for $x_1 = 0$)

$$-\gamma x_1=\tan^{-1}f-\phi_0\;, \tag{3.284}$$

and thus (recall that $x_3 = 0$ for $x_1 = 0$)

$$x_3=\frac{1}{\gamma}\log\frac{\cos\left(\gamma x_1-\phi_0\right)}{\cos\phi_0}\;. \tag{3.285}$$

b) For this case $\partial\log n/\partial x_3 = -1/x_3$, and thus Eq. (3.283) becomes [see Eq. (3.282)]

$$-\frac{1}{x_3}\left(1+\left(\frac{\mathrm{d}x_3}{\mathrm{d}x_1}\right)^2\right)=\frac{\mathrm{d}^2x_3}{\mathrm{d}x_1^2}\;. \tag{3.286}$$

The following holds

$$\frac{\mathrm{d}^2 x_3^2}{\mathrm{d}x_1^2} = \frac{\mathrm{d}}{\mathrm{d}x_1}\left(2x_3\frac{\mathrm{d}x_3}{\mathrm{d}x_1}\right) = 2\left(\left(\frac{\mathrm{d}x_3}{\mathrm{d}x_1}\right)^2 + x_3\frac{\mathrm{d}^2 x_3}{\mathrm{d}x_1^2}\right) , \tag{3.287}$$

and thus Eq. (3.286) can be rewritten as

$$\frac{\mathrm{d}^2 x_3^2}{\mathrm{d}x_1^2} = -2 . \tag{3.288}$$

The initial conditions lead to

$$x_1^2 + x_3^2 = L^2 . \tag{3.289}$$

8. According to the Fermat's principle, the trajectory can be found by solving the ray equation (3.42) for the case where the refractive index n is given by $n = \sqrt{Z/z}$ (note that conservation of the total energy $U+mv^2/2$ implies that the velocity v is related to the coordinate z by $v = \sqrt{-2gz}$). For motion in the xz plane the ray equation can be expressed as [see Eq. (3.283)]

$$\frac{\mathrm{d}\log n}{\mathrm{d}z}\left(1+f^2\right) = \frac{\mathrm{d}f}{\mathrm{d}x} , \tag{3.290}$$

where $\mathrm{d}\log n/\mathrm{d}z = -1/(2z)$ and where $f = \mathrm{d}z/\mathrm{d}x$. In terms of the angle α, which is related to f by

$$\alpha = \tan^{-1}\left(-\frac{1}{f}\right) , \tag{3.291}$$

one finds that

$$\begin{aligned}
\frac{\mathrm{d}(n\sin\alpha)}{\mathrm{d}z} &= n\frac{\mathrm{d}\sin\alpha}{\mathrm{d}z} + \sin\alpha\frac{\mathrm{d}n}{\mathrm{d}z} \\
&= n\left(\frac{\mathrm{d}\sin\left(\tan^{-1}\left(-\frac{1}{f}\right)\right)}{\mathrm{d}f}\frac{\frac{\mathrm{d}f}{\mathrm{d}x}}{\frac{\mathrm{d}z}{\mathrm{d}x}} + \sin\alpha\frac{\frac{\mathrm{d}n}{\mathrm{d}z}}{n}\right) \\
&= \frac{n}{(1+f^2)^{3/2}}\left(\frac{\mathrm{d}f}{\mathrm{d}x} - \frac{\mathrm{d}\log n}{\mathrm{d}z}\left(1+f^2\right)\right) ,
\end{aligned} \tag{3.292}$$

and thus [see Eq. (3.290)]

$$\frac{\mathrm{d}(n\sin\alpha)}{\mathrm{d}z} = 0 , \tag{3.293}$$

i.e. the trajectory locally satisfies Snell's law [compare with Eq. (3.55)]. With the help of Eqs. (3.292) and (3.293) one obtains (recall that $\mathrm{d}\log n/\mathrm{d}z = -1/(2z)$ and note that $f = -\cot\alpha$)

$$0 = n\left(\cos\alpha\frac{\mathrm{d}\alpha}{\mathrm{d}z} - \frac{\sin\alpha}{2z}\right) ,$$

and thus

$$\frac{\mathrm{d}z}{\mathrm{d}\alpha} = 2z\cot\alpha , \tag{3.294}$$

$$\frac{\mathrm{d}x}{\mathrm{d}\alpha} = \frac{\mathrm{d}x}{\mathrm{d}z}\frac{\mathrm{d}z}{\mathrm{d}\alpha} = -2z . \tag{3.295}$$

Integration leads to

$$z(\alpha) = R(-1+\cos 2\alpha) , \tag{3.296}$$

$$x(\alpha) = R(2\alpha - \sin 2\alpha) , \tag{3.297}$$

i.e. the trajectory which minimizes the travel time is a cycloid, and the constant R is its radius. The coordinate x can be expressed as a function of the coordinate z as

$$x = R\left(\cos^{-1}\frac{R+z}{R} - \sin\left(\cos^{-1}\frac{R+z}{R}\right)\right) . \tag{3.298}$$

The radius R is determined by solving

$$X = R\left(\cos^{-1}\frac{R+Z}{R} - \sin\left(\cos^{-1}\frac{R+Z}{R}\right)\right) . \tag{3.299}$$

9. The x_1 component of the ray equation (3.61) reads

$$\frac{\mathrm{d}}{\mathrm{d}s}\left(n\frac{\mathrm{d}x_1}{\mathrm{d}s}\right) = 0 . \tag{3.300}$$

The following holds [since s is an arc-length parameter, see Eq. (3.100)]

$$1 = \left(\frac{\mathrm{d}x_1}{\mathrm{d}s}\right)^2 + \left(\frac{\mathrm{d}x_3}{\mathrm{d}s}\right)^2 , \tag{3.301}$$

hence

$$\frac{\mathrm{d}}{\mathrm{d}s} = \frac{\mathrm{d}x_3}{\mathrm{d}s}\frac{\mathrm{d}}{\mathrm{d}x_3} = \frac{1}{\sqrt{1+\tan^2\theta}}\frac{\mathrm{d}}{\mathrm{d}x_3} = \cos\theta\frac{\mathrm{d}}{\mathrm{d}x_3} , \tag{3.302}$$

where

$$\tan\theta = \frac{\mathrm{d}x_1}{\mathrm{d}x_3} , \tag{3.303}$$

and thus Eq. (3.300), which can be rewritten as

$$\frac{\mathrm{d}}{\mathrm{d}x_3}(n\sin\theta) = 0 . \tag{3.304}$$

represents the Snell's law, according to which the term $n\sin\theta \equiv \alpha$ is a constant along the ray [compare with Eq. (3.55)]. Hence the following holds [see Eq. (3.303) and recall that $\tan\theta = \sin\theta/\sqrt{1-\sin^2\theta}$]

$$\frac{\mathrm{d}x_1}{\mathrm{d}x_3} = \frac{1}{\sqrt{\left(\frac{n}{\alpha}\right)^2 - 1}} . \tag{3.305}$$

By integration one finds that (recall that the ray passes through the origin point)

$$x_1 = \int_0^{x_3} \frac{\mathrm{d}x_3'}{\sqrt{\left(\frac{n(x_3')}{\alpha}\right)^2 - 1}} . \tag{3.306}$$

Alternatively, by using Eq. (3.283), which reads

$$\frac{\mathrm{d}\log n}{\mathrm{d}x_3}\left(1+f^2\right) = \frac{\mathrm{d}f}{\mathrm{d}x_1} , \tag{3.307}$$

where $f = \mathrm{d}x_3/\mathrm{d}x_1$ [see Eq. (3.282)], together with the relation (3.304), which implies that $\alpha = n\sin\theta = n/\sqrt{1+\cot^2\theta} = n/\sqrt{1+f^2}$ is a constant along the ray [see Eq. (3.303)], one finds that [the term n is replaced by $\alpha\sqrt{1+f^2}$, and the term $1+f^2$ is replaces by $(n/\alpha)^2$]

$$\frac{f}{1+f^2}\frac{\mathrm{d}f}{\mathrm{d}x_3}\left(\frac{n}{\alpha}\right)^2 = \frac{\mathrm{d}f}{\mathrm{d}x_1} , \tag{3.308}$$

hence

$$\frac{1}{f} = \frac{1}{\sqrt{\left(\frac{n}{\alpha}\right)^2 - 1}} , \tag{3.309}$$

in agreement with Eq. (3.305).

10. Consider an infinitesimal rope section of length $\mathrm{d}s$ located at the point $\mathbf{r} = (x_1, x_2, x_3)$. The contribution of this section to the total potential energy is $gx_3\lambda\mathrm{d}s$. The rope problem is mathematically equivalent to the problem of finding an optical ray that minimizes the time of flight in a medium having refractive index given by $n = x_3/x_0$, where x_0 is a constant. According to Fermat's principle, the solution to the optical minimization problem satisfies the ray equation (3.61). The solution given by Eq. (3.306) yields (it is assumed that the ray lays in the $x_2 = 0$ plane)

$$\frac{x_1}{x_0} = \frac{1}{x_0}\int \frac{\mathrm{d}x_3'}{\sqrt{\left(\frac{x_3'}{x_0}\right)^2 - 1}} = \frac{x_{10}}{x_0} + \cosh^{-1}\frac{x_3}{x_0} , \tag{3.310}$$

where x_{10} is a constant, thus

$$\frac{x_3}{x_0} = \cosh\left(\frac{x_1 - x_{10}}{x_0}\right) . \tag{3.311}$$

This curve is commonly called the catenary curve. The constants x_0 and x_{10} are determined by the locations of the clamping points and by the total length of the rope.

11. Let $\mathbf{x}_{\mathrm{C}}(q) = (r(q), \theta(q), \phi(q))$ be a general parametrization of the optical ray. For the case of spherical coordinates the following holds

$$\dot{\mathbf{x}}_{\mathrm{C}}^2 = \dot{r}^2 + \left(r\dot{\theta}\right)^2 + \left(r \sin\theta \dot{\phi}\right)^2 , \tag{3.312}$$

where overdot denotes a derivative with respect to the parameter q, and thus the ray equation corresponding to the coordinate ϕ in arc-length parametrization for the case of spherically symmetric medium is given by [compare with Eqs. (3.111), (3.112) and (3.113)]

$$0 = \frac{\mathrm{d}}{\mathrm{d}s}\left(nr^2 \sin^2\theta \frac{\mathrm{d}\phi}{\mathrm{d}s}\right) . \tag{3.313}$$

The above result (3.313) implies that $\mathrm{d}L_3/\mathrm{d}s = 0$ [see Eq. (3.211)]. Recall Eq. (3.75), which states that for the case of a spherically symmetric medium the following holds

$$\frac{\mathrm{d}}{\mathrm{d}s}(n\mathbf{r} \times \hat{\mathbf{s}}) = 0 . \tag{3.314}$$

With the help of the general vector identity

$$(\mathbf{A} \times \mathbf{B}) \cdot (\mathbf{C} \times \mathbf{D}) = (\mathbf{A} \cdot \mathbf{C})(\mathbf{B} \cdot \mathbf{D}) - (\mathbf{A} \cdot \mathbf{D})(\mathbf{B} \cdot \mathbf{C}) , \tag{3.315}$$

one finds that [see Eqs. (3.42) and (3.210)]

$$\begin{aligned}(n\mathbf{r} \times \hat{\mathbf{s}}) \cdot (n\mathbf{r} \times \hat{\mathbf{s}}) &= n^2 r^2 \left(\frac{\mathbf{r}}{r} \times \hat{\mathbf{s}}\right) \cdot \left(\frac{\mathbf{r}}{r} \times \hat{\mathbf{s}}\right) \\ &= n^2 r^2 \left[1 - \left(\frac{\mathbf{r}}{r} \cdot \frac{\mathrm{d}\mathbf{r}}{\mathrm{d}s}\right)^2\right] \\ &= n^2 r^2 \left[1 - \left(\frac{\mathrm{d}r}{\mathrm{d}s}\right)^2\right] \\ &= L^2 ,\end{aligned} \tag{3.316}$$

and thus $\mathrm{d}L^2/\mathrm{d}s = 0$.

12. In a spherically symmetric medium the vector $n\mathbf{r}\times\hat{\mathbf{s}}$ is a constant along an optical ray $\mathbf{r}(s)$ [see Eq. (3.75)]. Thus, the optical ray is a curve in a plane containing the origin. In spherical coordinates (r,θ,ϕ) the plane is taken to be a plane of a constant angle ϕ. For this case an optical ray in arc-length parametrization satisfies [see Eq. (3.312)]

$$1=\left(\frac{\mathrm{d}r}{\mathrm{d}s}\right)^2+\left(r\frac{\mathrm{d}\theta}{\mathrm{d}s}\right)^2 . \tag{3.317}$$

The variable L^2 given by Eq. (3.210) is a constant along an optical ray in a spherically symmetric medium, and thus

$$\frac{\mathrm{d}r}{\mathrm{d}s}=\sqrt{1-\frac{L^2}{n^2r^2}} , \tag{3.318}$$

and [see Eq. (3.317)]

$$\frac{\mathrm{d}\theta}{\mathrm{d}s}=\frac{L}{nr^2} , \tag{3.319}$$

and thus

$$\frac{\mathrm{d}\theta}{\mathrm{d}r}=\frac{L}{r\sqrt{n^2r^2-L^2}} . \tag{3.320}$$

or

$$\frac{\mathrm{d}\theta}{\mathrm{d}\rho}=\frac{1}{\rho\sqrt{\left(\frac{n\rho}{l}\right)^2-1}} , \tag{3.321}$$

where

$$\rho=\frac{r}{R} , \tag{3.322}$$

$$l=\frac{L}{R} . \tag{3.323}$$

a) For this case Eq. (3.321) yields [see Eq. (3.213)]

$$\frac{\mathrm{d}\theta}{\mathrm{d}\rho}=\frac{1}{\rho\sqrt{\frac{(2-\rho^2)\rho^2}{l^2}-1}} . \tag{3.324}$$

The variable transformation

$$\rho=\frac{l}{\sqrt{1+q}} \tag{3.325}$$

leads to

$$\frac{\mathrm{d}\theta}{\mathrm{d}q}=-\frac{1}{2\sqrt{1-l^2-q^2}} , \tag{3.326}$$

and the transformation

$$q = \sqrt{1 - l^2} \cos \psi \; , \tag{3.327}$$

to

$$\frac{\mathrm{d}\theta}{\mathrm{d}\psi} = \frac{1}{2} \; , \tag{3.328}$$

and thus

$$\cos(2\theta + 2\psi_0) = \frac{\left(\frac{L}{r}\right)^2 - 1}{\sqrt{1 - \left(\frac{L}{R}\right)^2}} \; , \tag{3.329}$$

where ψ_0 is a constant, which can be taken to be zero. Using the notation

$$r_0^2 = \frac{L^2}{1 - \sqrt{1 - \left(\frac{L}{R}\right)^2}} \; , \tag{3.330}$$

$$\eta = \sqrt{1 - \frac{L^2}{R^2}} \; , \tag{3.331}$$

Eq. (3.329) can be rewritten as

$$\left(\frac{r}{r_0}\right)^2 = \frac{1 - \eta}{1 + \eta \cos 2\theta} \; , \tag{3.332}$$

and thus the optical ray is an ellipse having eccentricity e given by [compare with Eq. (2.162)]

$$e = \sqrt{\frac{2\eta}{1 + \eta}} = \frac{2\sqrt{1 - \left(\frac{L}{R}\right)^2}}{1 + \sqrt{1 - \left(\frac{L}{R}\right)^2}} \; . \tag{3.333}$$

b) For this case Eq. (3.321) yields [see Eq. (3.214)]

$$\frac{\mathrm{d}\theta}{\mathrm{d}\rho} = \frac{1}{\rho\sqrt{\left(\frac{n_0 \rho}{l(1+\rho^2)}\right)^2 - 1}} \; . \tag{3.334}$$

Integration leads to

$$\theta - \theta_0 = \arcsin\left(\frac{1}{\sqrt{\left(\frac{n_0}{l}\right)^2 - 4}} \frac{\rho^2 - 1}{\rho}\right) \; , \tag{3.335}$$

where θ_0 is a constant, and thus

$$r\sin(\theta-\theta_0)=\frac{r^2-R^2}{R\sqrt{\left(\frac{n_0R}{L}\right)^2-4}}\;. \tag{3.336}$$

In Cartesian coordinates

$$x=r\cos\theta\;, \tag{3.337}$$

$$y=r\sin\theta\;, \tag{3.338}$$

Eq. (3.336) becomes

$$y\cos\theta_0-x\sin\theta_0=\frac{x^2+y^2-R^2}{R\sqrt{\left(\frac{n_0R}{L}\right)^2-4}}\;, \tag{3.339}$$

or

$$x_0^2+y_0^2+R^2=(x+x_0)^2+(y-y_0)^2\;, \tag{3.340}$$

where

$$x_0=R\sqrt{\left(\frac{n_0R}{2L}\right)^2-1}\sin\theta_0\;, \tag{3.341}$$

$$y_0=R\sqrt{\left(\frac{n_0R}{2L}\right)^2-1}\cos\theta_0\;, \tag{3.342}$$

i.e. the optical ray is a circle centered at $(-x_0,y_0)$ having radius

$$\sqrt{x_0^2+y_0^2+R^2}=\frac{n_0R^2}{2L}\;. \tag{3.343}$$

13. For the case of a spherically symmetric medium the optical rays in spherical coordinates can be evaluated by solving Eq. (3.320), which for the current case becomes

$$\frac{\mathrm{d}\theta}{\mathrm{d}r}=\frac{L}{r\sqrt{\left(1+\frac{r_\mathrm{S}}{r}\right)^2r^2-L^2}}\;, \tag{3.344}$$

where L is a constant along an optical ray, and where the so-called Schwarzschild radius r_S is given by

$$r_\mathrm{S}=\frac{2GM}{c^2}\;. \tag{3.345}$$

Integration yields

$$\begin{aligned}\theta_\mathrm{r}&=\int\frac{L}{r\sqrt{\left(1+\frac{r_\mathrm{S}}{r}\right)^2r^2-L^2}}\mathrm{d}r\\&=\frac{L}{\sqrt{L^2-r_\mathrm{S}^2}}\tan^{-1}\frac{\rho r_\mathrm{S}-L^2}{\sqrt{L^2-r_\mathrm{S}^2}\sqrt{\rho^2-L^2}}\;,\end{aligned} \tag{3.346}$$

where $\theta_{\mathrm{r}} = \theta - \theta_0$, the angle θ_0 is a constant and $\rho = r_{\mathrm{S}} + r$. The following holds [see Eq. (3.346)]

$$\lim_{\rho \to L} \theta_{\mathrm{r}} = \frac{\pi L}{2\sqrt{L^2 - r_{\mathrm{S}}^2}} = \frac{\pi}{2} + O\left(r_{\mathrm{S}}^2\right) , \tag{3.347}$$

and

$$\lim_{\rho \to \infty} \theta_{\mathrm{r}} = \frac{L}{\sqrt{L^2 - r_{\mathrm{S}}^2}} \tan^{-1} \frac{r_{\mathrm{S}}}{\sqrt{L^2 - r_{\mathrm{S}}^2}} = \frac{r_{\mathrm{S}}}{L} + O\left(r_{\mathrm{S}}^3\right) . \tag{3.348}$$

The smallest distance r_0 between the optical ray and the point mass (for which $\rho = L$) is given by $r_0 = L - r_{\mathrm{S}}$, and thus the deflection angle α is given by

$$\alpha = 2\frac{r_{\mathrm{S}}}{r_0} + O\left(r_{\mathrm{S}}^2\right) . \tag{3.349}$$

14. With the help of the aberration of light formula (1.62) one finds that

$$\cos\theta_{\mathrm{i}}' = \frac{\cos\theta + \beta}{1 + \beta\cos\theta} , \tag{3.350}$$

$$\cos\theta_{\mathrm{r}}' = \frac{\cos\theta - \beta}{1 - \beta\cos\theta} . \tag{3.351}$$

a) The above results (3.350) and (3.351) yield

$$\cos\theta_{\mathrm{r}}' = \frac{\left(1 + \beta^2\right)\cos\theta_{\mathrm{i}}' - 2\beta}{1 - 2\beta\cos\theta_{\mathrm{i}}' + \beta^2} . \tag{3.352}$$

Alternatively, with the help of Eqs. (3.350) and (3.351) one finds that

$$\frac{\sin\theta_{\mathrm{i}}' + \sin\theta_{\mathrm{r}}'}{\sin\left(\theta_{\mathrm{r}}' - \theta_{\mathrm{i}}'\right)} = \frac{\sin\theta_{\mathrm{i}}' + \sin\theta_{\mathrm{r}}'}{\sin\theta_{\mathrm{r}}'\cos\theta_{\mathrm{i}}' - \cos\theta_{\mathrm{r}}'\sin\theta_{\mathrm{i}}'} = \beta . \tag{3.353}$$

b) Consider the case where the reflecting plane of the moving mirror at time t is the plane $z = -\beta ct$. Consider a reflection process from the initial point $\mathbf{r}_{\mathrm{A}} = (-\alpha L, 0, L)$ to the final point $\mathbf{r}_{\mathrm{B}} = (\alpha L, 0, L)$, where $\alpha > 0$ is dimensionless and $L > 0$ is the distance between both points $\mathbf{r}_{\mathrm{A}}$ and $\mathbf{r}_{\mathrm{B}}$ and the mirror at time $t = 0$. The optical ray is made of two straight sections, the first from the point $\mathbf{r}_{\mathrm{A}}$ at time $t = 0$ to a point on the mirror $\mathbf{r}_{\mathrm{M}} = (\eta L, 0, -\beta c t_1)$ at time $t_1 > 0$, where η is dimensionless, and the second is from the point $\mathbf{r}_{\mathrm{M}}$ at time t_1 to the point $\mathbf{r}_{\mathrm{B}}$ at some later time. The following holds

$$\begin{aligned}\frac{\mathbf{r}_{\mathrm{A}} - \mathbf{r}_{\mathrm{M}}}{|\mathbf{r}_{\mathrm{A}} - \mathbf{r}_{\mathrm{M}}|} &= \frac{\left(-\left(\alpha + \eta\right)L, 0, L + \beta c t_1\right)}{\sqrt{\left(\alpha + \eta\right)^2 L^2 + \left(L + \beta c t_1\right)^2}} \\ &= \left(\sin\theta_{\mathrm{i}}', 0, \cos\theta_{\mathrm{i}}'\right) ,\end{aligned} \tag{3.354}$$

$$\frac{\mathbf{r}_{\mathrm{B}}-\mathbf{r}_{\mathrm{M}}}{|\mathbf{r}_{\mathrm{B}}-\mathbf{r}_{\mathrm{M}}|} = \frac{((\alpha-\eta)L, 0, L+\beta ct_1)}{\sqrt{(\alpha-\eta)^2 L^2 + (L+\beta ct_1)^2}} = \left(\sin\theta'_{\mathrm{r}}, 0, \cos\theta'_{\mathrm{r}}\right) \ . \tag{3.355}$$

The requirement that the speed of light is c yields

$$ct_1 = |\mathbf{r}_{\mathrm{A}}-\mathbf{r}_{\mathrm{M}}| = \sqrt{(\alpha+\eta)^2 L^2 + (L+\beta ct_1)^2} \ . \tag{3.356}$$

This condition (3.356) can be rewritten as

$$g\left(\eta, \frac{ct_1}{L}\right) = 0 \ , \tag{3.357}$$

where

$$g(x,z) = (\alpha+x)^2 + (1+\beta z)^2 - z^2 \ .$$

The total time of flight T can be expressed as

$$T = \frac{L}{c}\tau\left(\eta, \frac{ct_1}{L}\right) \ , \tag{3.358}$$

where

$$\tau(x,z) = \left(z + \sqrt{(\alpha-x)^2 + (1+\beta z)^2}\right) \ . \tag{3.359}$$

For the optical trajectory that minimizes T the following holds

$$\boldsymbol{\nabla}\tau = \xi\boldsymbol{\nabla}g \ , \tag{3.360}$$

where $\boldsymbol{\nabla} = (\partial/\partial x, \partial/\partial z)$ is a two-dimensional gradient and ξ is a Lagrange multiplier. Condition (3.360), which can be rewritten as

$$\frac{\frac{\partial\tau}{\partial x}}{\frac{\partial g}{\partial x}} = \frac{\frac{\partial\tau}{\partial z}}{\frac{\partial g}{\partial z}} \ , \tag{3.361}$$

or

$$\frac{-\frac{(\alpha-x)}{\sqrt{(\alpha-x)^2+(1+\beta z)^2}}}{2(\alpha+x)} = \frac{1+\frac{\beta(1+\beta z)}{\sqrt{(\alpha-x)^2+(1+\beta z)^2}}}{2\beta(1+\beta z)-2z} \ , \tag{3.362}$$

together with Eqs. (3.354) and (3.355) lead to

$$\frac{\sin\theta'_{\mathrm{r}}}{\sin\theta'_{\mathrm{i}}} = \frac{1+\beta\cos\theta'_{\mathrm{r}}}{\beta\cos\theta'_{\mathrm{i}}-1} \ , \tag{3.363}$$

and thus

$$\frac{\sin\theta'_{\mathrm{i}}+\sin\theta'_{\mathrm{r}}}{\sin\left(\theta'_{\mathrm{r}}-\theta'_{\mathrm{i}}\right)} = \beta \ , \tag{3.364}$$

in agreement with Eq. (3.353).

15. Using Eqs. (1.96) and (3.41) one finds that

$$-\mu\boldsymbol{\nabla}\times\mathbf{H}_0-\boldsymbol{\nabla}\psi\times(\boldsymbol{\nabla}\times\mathbf{E}_0)=n^2\left(\hat{\mathbf{s}}\cdot\mathbf{E}_1\right)\hat{\mathbf{s}}\;, \tag{3.365}$$

or [see Eq. (3.35)]

$$-\mu\left[\boldsymbol{\nabla}\times\left[\frac{\boldsymbol{\nabla}\psi}{\mu}\times\mathbf{E}_0\right]+\frac{\boldsymbol{\nabla}\psi}{\mu}\times(\boldsymbol{\nabla}\times\mathbf{E}_0)\right]=n^2\left(\hat{\mathbf{s}}\cdot\mathbf{E}_1\right)\hat{\mathbf{s}}\;. \tag{3.366}$$

Solution existence of the above equation requires that the vector $\mathbf{F}_0$, which is defined as

$$\mathbf{F}_0\equiv\boldsymbol{\nabla}\times\left[\frac{\boldsymbol{\nabla}\psi}{\mu}\times\mathbf{E}_0\right]+\frac{\boldsymbol{\nabla}\psi}{\mu}\times(\boldsymbol{\nabla}\times\mathbf{E}_0)\;, \tag{3.367}$$

is parallel to $\hat{\mathbf{s}}$. In what follows we rewrite $\mathbf{F}_0$ in a way allowing identifying the components parallel and orthogonal to $\hat{\mathbf{s}}$. Using Eq. (3.118) one finds that

$$\begin{aligned}\mathbf{F}_0=&\frac{\boldsymbol{\nabla}\psi}{\mu}(\boldsymbol{\nabla}\cdot\mathbf{E}_0)-\mathbf{E}_0\left(\boldsymbol{\nabla}\cdot\frac{\boldsymbol{\nabla}\psi}{\mu}\right)+\boldsymbol{\nabla}\left(\frac{\boldsymbol{\nabla}\psi}{\mu}\cdot\mathbf{E}_0\right)\\&-\mathbf{E}_0\times\left(\boldsymbol{\nabla}\times\frac{\boldsymbol{\nabla}\psi}{\mu}\right)-2\left(\frac{\boldsymbol{\nabla}\psi}{\mu}\cdot\boldsymbol{\nabla}\right)\mathbf{E}_0\;.\end{aligned} \tag{3.368}$$

The third term on the right vanishes [see Eq. (3.36)]. Moreover, using Eq. (2.150) and $\boldsymbol{\nabla}\times\boldsymbol{\nabla}\psi=0$ one finds that

$$\begin{aligned}\mathbf{F}_0=&\frac{\boldsymbol{\nabla}\psi}{\mu}(\boldsymbol{\nabla}\cdot\mathbf{E}_0)-\mathbf{E}_0\left(\boldsymbol{\nabla}\cdot\frac{\boldsymbol{\nabla}\psi}{\mu}\right)\\&-\mathbf{E}_0\times\left[\boldsymbol{\nabla}\left(\frac{1}{\mu}\right)\times\boldsymbol{\nabla}\psi\right]-2\left(\frac{\boldsymbol{\nabla}\psi}{\mu}\cdot\boldsymbol{\nabla}\right)\mathbf{E}_0\;.\end{aligned} \tag{3.369}$$

In addition, using the vector identity (1.96), which is given by

$$\mathbf{A}\times(\mathbf{B}\times\mathbf{C})=(\mathbf{A}\cdot\mathbf{C})\,\mathbf{B}-(\mathbf{A}\cdot\mathbf{B})\,\mathbf{C}\;, \tag{3.370}$$

for the third term on the right and using again Eq. (3.36) leads to

$$\begin{aligned}\mathbf{F}_0=&\frac{\boldsymbol{\nabla}\psi}{\mu}(\boldsymbol{\nabla}\cdot\mathbf{E}_0)-\mathbf{E}_0\left(\boldsymbol{\nabla}\cdot\frac{\boldsymbol{\nabla}\psi}{\mu}\right)\\&+\left(\mathbf{E}_0\cdot\boldsymbol{\nabla}\left(\frac{1}{\mu}\right)\right)\boldsymbol{\nabla}\psi-2\left(\frac{\boldsymbol{\nabla}\psi}{\mu}\cdot\boldsymbol{\nabla}\right)\mathbf{E}_0\;.\end{aligned} \tag{3.371}$$

Using Eq. (2.149) for the second term on the right and multiplying by μ leads to

$$\mu \mathbf{F}_0 = \boldsymbol{\nabla}\psi \left(\boldsymbol{\nabla}\cdot \mathbf{E}_0\right) - \mathbf{E}_0 \left[\boldsymbol{\nabla}^2\psi - \boldsymbol{\nabla}\left(\log \mu\right)\cdot \boldsymbol{\nabla}\psi\right] \\ - \left[\mathbf{E}_0 \cdot \boldsymbol{\nabla}\left(\log \mu\right)\right] \boldsymbol{\nabla}\psi - 2\left(\boldsymbol{\nabla}\psi\cdot\boldsymbol{\nabla}\right)\mathbf{E}_0 \; . \tag{3.372}$$

The forth term on the right can be rewritten using Eqs. (3.44) and (3.154) as

$$\begin{aligned}\left(\boldsymbol{\nabla}\psi\cdot\boldsymbol{\nabla}\right)\mathbf{E}_0 &= n\frac{\mathrm{d}}{\mathrm{d}s}\left(\alpha_{\nu 0}\hat{\boldsymbol{\nu}} + \alpha_{b0}\hat{\mathbf{b}}\right) \\ &= n\left[\frac{\mathrm{d}\alpha_{\nu 0}}{\mathrm{d}s}\hat{\boldsymbol{\nu}} + \frac{\mathrm{d}\alpha_{b0}}{\mathrm{d}s}\hat{\mathbf{b}} + \alpha_{\nu 0}\left(-\kappa\hat{\mathbf{s}} + \tau\hat{\mathbf{b}}\right) - \alpha_{b0}\tau\hat{\boldsymbol{\nu}}\right] \\ & n\left[\left(\frac{\mathrm{d}\alpha_{\nu 0}}{\mathrm{d}s} - \alpha_{b0}\tau\right)\hat{\boldsymbol{\nu}} + \left(\frac{\mathrm{d}\alpha_{b0}}{\mathrm{d}s} + \alpha_{\nu 0}\tau\right)\hat{\mathbf{b}} - \alpha_{\nu 0}\kappa\hat{\mathbf{s}}\right] \; ,\end{aligned} \tag{3.373}$$

thus, using Eqs. (3.36) and (3.56) one finds that

$$\hat{\mathbf{s}}\cdot\left[\left(\boldsymbol{\nabla}\psi\cdot\boldsymbol{\nabla}\right)\mathbf{E}_0\right] = -n\mathbf{E}_0\cdot\kappa\hat{\boldsymbol{\nu}} = -\,\mathbf{E}_0\cdot\boldsymbol{\nabla}n \; . \tag{3.374}$$

Using the last result together with Eq. (3.372) one can write the condition for $\mathbf{F}_0$ to be parallel to $\hat{\mathbf{s}}$ as

$$0 = -\mathbf{E}_0\left[\boldsymbol{\nabla}^2\psi - \boldsymbol{\nabla}\left(\log\mu\right)\cdot\boldsymbol{\nabla}\psi\right] - 2\left(\boldsymbol{\nabla}\psi\cdot\boldsymbol{\nabla}\right)\mathbf{E}_0 - 2\hat{\mathbf{s}}\left(\mathbf{E}_0\cdot\boldsymbol{\nabla}n\right) \; , \tag{3.375}$$

or

$$2\left(\boldsymbol{\nabla}\psi\cdot\boldsymbol{\nabla}\right)\mathbf{E}_0 + \mathbf{E}_0\left[\boldsymbol{\nabla}^2\psi - \boldsymbol{\nabla}\left(\log\mu\right)\cdot\boldsymbol{\nabla}\psi\right] + 2\left[\mathbf{E}_0\cdot\boldsymbol{\nabla}\left(\log n\right)\right]\boldsymbol{\nabla}\psi = 0 \; , \tag{3.376}$$

in agreement with Eq. (3.114).

16. By using Eq. (2.150) and $\psi = n\hat{\mathbf{s}}$ one finds that

$$0 = \boldsymbol{\nabla}\times\boldsymbol{\nabla}\psi = \boldsymbol{\nabla}\times\left(n\hat{\mathbf{s}}\right) = n\boldsymbol{\nabla}\times\hat{\mathbf{s}} + \left(\boldsymbol{\nabla}n\right)\times\hat{\mathbf{s}} \; , \tag{3.377}$$

thus, by multiplying by $\hat{\mathbf{s}}$ one obtains

$$\hat{\mathbf{s}}\cdot\boldsymbol{\nabla}\times\hat{\mathbf{s}} = 0 \; , \tag{3.378}$$

and

$$\boldsymbol{\nabla}\times\hat{\mathbf{s}} = \hat{\mathbf{s}}\times\left(\boldsymbol{\nabla}\log n\right) \; . \tag{3.379}$$

Therefor, the following hold [see Eq. (3.56)]

$$\hat{\boldsymbol{\nu}}\cdot\boldsymbol{\nabla}\times\hat{\mathbf{s}} = \hat{\boldsymbol{\nu}}\cdot\left[\hat{\mathbf{s}}\times\left(\boldsymbol{\nabla}\log n\right)\right] = -\left(\boldsymbol{\nabla}\log n\right)\cdot\hat{\mathbf{b}} = 0 \; , \tag{3.380}$$

and

$$\hat{\mathbf{b}}\cdot\boldsymbol{\nabla}\times\hat{\mathbf{s}} = \hat{\mathbf{b}}\cdot\left[\hat{\mathbf{s}}\times\left(\boldsymbol{\nabla}\log n\right)\right] = \left(\boldsymbol{\nabla}\log n\right)\cdot\hat{\boldsymbol{\nu}} = \kappa \; . \tag{3.381}$$

17. By multiplying Eq. (3.114) by $\mathbf{E}_0$ one obtains

$$2\mathbf{E}_0\cdot\left(\boldsymbol{\nabla}\psi\cdot\boldsymbol{\nabla}\right)\mathbf{E}_0+\mathbf{E}_0^2\left[\boldsymbol{\nabla}^2\psi-\boldsymbol{\nabla}\left(\log\mu\right)\cdot\boldsymbol{\nabla}\psi\right]+2\left[\mathbf{E}_0\cdot\boldsymbol{\nabla}\left(\log n\right)\right]\left(\mathbf{E}_0\cdot\boldsymbol{\nabla}\psi\right)=0\;. \tag{3.382}$$

As can be seen from Eq. (3.36), the last term vanishes. Using Eq. (2.149) one finds that

$$\boldsymbol{\nabla}^2\psi-\boldsymbol{\nabla}\left(\log\mu\right)\cdot\boldsymbol{\nabla}\psi=\mu\boldsymbol{\nabla}\cdot\left(\frac{\boldsymbol{\nabla}\psi}{\mu}\right)\;, \tag{3.383}$$

thus

$$2\mathbf{E}_0\cdot\left(\frac{\boldsymbol{\nabla}\psi}{\mu}\cdot\boldsymbol{\nabla}\right)\mathbf{E}_0+\mathbf{E}_0^2\boldsymbol{\nabla}\cdot\left(\frac{\boldsymbol{\nabla}\psi}{\mu}\right)=0\;. \tag{3.384}$$

Using $\boldsymbol{\nabla}\psi=n\hat{\mathbf{s}}$ and $n=\sqrt{\epsilon\mu}$ lead to

$$2\sqrt{\frac{\epsilon}{\mu}}\mathbf{E}_0\cdot\left(\hat{\mathbf{s}}\cdot\boldsymbol{\nabla}\right)\mathbf{E}_0+\mathbf{E}_0^2\boldsymbol{\nabla}\cdot\left(\sqrt{\frac{\epsilon}{\mu}}\hat{\mathbf{s}}\right)=0\;, \tag{3.385}$$

or

$$\sqrt{\frac{\epsilon}{\mu}}\hat{\mathbf{s}}\cdot\boldsymbol{\nabla}\mathbf{E}_0^2+\mathbf{E}_0^2\boldsymbol{\nabla}\cdot\left(\sqrt{\frac{\epsilon}{\mu}}\hat{\mathbf{s}}\right)=0\;, \tag{3.386}$$

thus [see Eq. (2.149)]

$$\boldsymbol{\nabla}\cdot\left(\mathbf{E}_0^2\sqrt{\frac{\epsilon}{\mu}}\hat{\mathbf{s}}\right)=0\;. \tag{3.387}$$

Define the ratio [see Eq. (2.165)]

$$\eta\equiv\mathbf{E}_0^2/\left|\mathbf{E}_0\right|^2\;. \tag{3.388}$$

According to Eqs. (3.166) and (3.387) both $\left|\mathbf{E}_0\right|^2$ and $\mathbf{E}_0^2$ have the same s dependence

$$\mathbf{E}_0^2\left(s\right)\sqrt{\frac{\epsilon\left(s\right)}{\mu\left(s\right)}}=\mathbf{E}_0^2\left(s_0\right)\sqrt{\frac{\epsilon\left(s_0\right)}{\mu\left(s_0\right)}}\exp\left[-\int_{s_0}^{s}\mathrm{d}s'\;\left(\boldsymbol{\nabla}\cdot\hat{\mathbf{s}}\right)\right]\;. \tag{3.389}$$

Thus, η 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. (2.162).

18. The ray equation is given by [see Eq. (3.61)]

$$\frac{\mathrm{d}}{\mathrm{d}s}\left(n\frac{\mathrm{d}\mathbf{r}}{\mathrm{d}s}\right)=\frac{\boldsymbol{\nabla}n^2}{2n}=-\frac{n_0^2K^2\left(x\hat{\mathbf{x}}+y\hat{\mathbf{y}}\right)}{n}\;. \tag{3.390}$$

Since $x^2(s) + y^2(s) = R^2$ the refractive index along the ray is a constant given by

$$n_{\mathrm{r}} = n_0 \sqrt{1 - K^2 R^2} \,, \tag{3.391}$$

and thus the ray equation (3.390) becomes

$$\frac{\mathrm{d}^2 \mathbf{r}}{\mathrm{d}s^2} = -k^2 \left(x \hat{\mathbf{x}} + y \hat{\mathbf{y}} \right) \,, \tag{3.392}$$

where

$$k = \frac{K}{\sqrt{1 - K^2 R^2}} \,. \tag{3.393}$$

The general solution for which $x^2(s) + y^2(s) = R^2$ is given by

$$x(s) = R \cos(ks + \phi_0) \,, \tag{3.394}$$

$$y(s) = R \sin(ks + \phi_0) \,, \tag{3.395}$$

$$z(s) = z_0 + \alpha k s \,, \tag{3.396}$$

where ϕ_0, z_0 and α are constants. The requirement that

$$1 = \left| \frac{\mathrm{d}\mathbf{r}}{\mathrm{d}s} \right| = \sqrt{k^2 \left(R^2 + \alpha^2 \right)} \,, \tag{3.397}$$

yields

$$\alpha = \frac{\sqrt{1 - k^2 R^2}}{k} \,. \tag{3.398}$$

As can be seen from Eqs. (3.142) and (3.144)

$$\frac{\mathrm{d}\phi_{\mathrm{e}}}{\mathrm{d}s} = \tau \,, \tag{3.399}$$

where τ is a torsion, which for the case of an helix is given by Eq. (3.258), thus

$$\frac{\mathrm{d}\phi_{\mathrm{e}}}{\mathrm{d}s} = \frac{\alpha}{R^2 + \alpha^2} = \frac{K\sqrt{1 - 2K^2 R^2}}{1 - K^2 R^2} , \tag{3.400}$$

where $2K^2 R^2 \leq 1$.

19. The action (3.222) can be expressed in terms of a Lagrangian $\mathcal{L}$ as [see Eq. (1.15)]

$$S = \int_{t_1}^{t_2} \mathrm{d}t \, \mathcal{L} \,, \tag{3.401}$$

where $\mathcal{L}$ is given by

$$\mathcal{L} = -mc^2\sqrt{1-\frac{\dot{\mathbf{r}}\cdot\dot{\mathbf{r}}}{c^2}} + \frac{q}{c}\dot{\mathbf{r}}\cdot\mathbf{A} - q\phi \ , \tag{3.402}$$

and where overdot denotes a derivative with respect to time t, i.e. $\dot{\mathbf{r}} = \mathrm{d}\mathbf{r}/\mathrm{d}t = (\mathrm{d}x_1/\mathrm{d}t, \mathrm{d}x_2/\mathrm{d}t, \mathrm{d}x_3/\mathrm{d}t)$. Consider a trajectory

$$\mathbf{r}(t) = \mathbf{r}_{\mathrm{c}}(t) + \delta\mathbf{r}(t) \ , \tag{3.403}$$

where $\mathbf{r}_{\mathrm{c}}(t)$ is assumed to be a classical trajectory and where $\delta\mathbf{r}(t)$ is considered as infinitesimally small. It is assume that $\mathbf{r}(t_1) = \mathbf{r}_{\mathrm{c}}(t_1) = \mathbf{r}_1$ and $\mathbf{r}(t_2) = \mathbf{r}_{\mathrm{c}}(t_2) = \mathbf{r}_2$, 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 δS is given by

$$\delta S = \int_{t_1}^{t_2} \mathrm{d}t\ \delta\mathcal{L} = \int_{t_1}^{t_2} \mathrm{d}t\ \sum_{n=1}^{3}\left(\frac{\partial\mathcal{L}}{\partial x_n}\delta x_n + \frac{\partial\mathcal{L}}{\partial\dot{x}_n}\frac{\mathrm{d}}{\mathrm{d}t}\delta x_n\right) . \tag{3.404}$$

Integrating the second term by parts leads to

$$\begin{aligned}\delta S = &\int_{t_1}^{t_2} \mathrm{d}t\ \sum_{n=1}^{3}\left(\frac{\partial\mathcal{L}}{\partial x_n} - \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial\mathcal{L}}{\partial\dot{x}_n}\right)\delta x_n \\ &+\sum_{n=1}^{N}\left[\frac{\partial\mathcal{L}}{\partial\dot{x}_n}\delta x_n\right|_{t_1}^{t_2} .\end{aligned} \tag{3.405}$$

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 δx_n, and thus

$$\frac{\partial\mathcal{L}}{\partial x_n} = \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial\mathcal{L}}{\partial\dot{x}_n} . \tag{3.406}$$

The set of equations (3.406) is called the Euler-Lagrange equations. For the coordinate x_1 Eq. (3.406) reads

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial\mathcal{L}}{\partial\dot{x}_1} = \frac{\partial\mathcal{L}}{\partial x_1} \ , \tag{3.407}$$

and the following holds

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial\mathcal{L}}{\partial\dot{x}_1} = \dot{p}_1 + \frac{q}{c}\left(\frac{\partial A_1}{\partial t} + \dot{x}_1\frac{\partial A_1}{\partial x_1} + \dot{x}_2\frac{\partial A_1}{\partial x_2} + \dot{x}_3\frac{\partial A_1}{\partial x_3}\right) \ , \tag{3.408}$$

where

$$p_1 = m\gamma\dot{x}_1 \ , \tag{3.409}$$

$$\gamma = \frac{1}{\sqrt{1 - \frac{\dot{\mathbf{r}}\cdot\dot{\mathbf{r}}}{c^2}}} , \tag{3.410}$$

and

$$\frac{\partial \mathcal{L}}{\partial x_1} = -q \frac{\partial \varphi}{\partial x_1} + \frac{q}{c}\left(\dot{x}_1 \frac{\partial A_1}{\partial x_1} + \dot{x}_2 \frac{\partial A_2}{\partial x} + \dot{x}_3 \frac{\partial A_3}{\partial x} \right) , \tag{3.411}$$

thus [see Eqs. (2.23) and (2.25)]

$$\begin{aligned} \dot{p}_1 &= \underbrace{-q \frac{\partial \varphi}{\partial x_1} - \frac{q}{c}\frac{\partial A_1}{\partial t}}_{qE_1} \\ &\quad + \frac{q}{c}\left[\underbrace{\dot{x}_2 \underbrace{\left(\frac{\partial A_2}{\partial x_1} - \frac{\partial A_1}{\partial x_2} \right)}_{(\boldsymbol{\nabla}\times\mathbf{A})_3} - \dot{x}_3 \underbrace{\left(\frac{\partial A_1}{\partial x_3} - \frac{\partial A_3}{\partial x_1} \right)}_{(\boldsymbol{\nabla}\times\mathbf{A})_2}}_{(\dot{\mathbf{r}}\times(\boldsymbol{\nabla}\times\mathbf{A}))_1} \right] , \end{aligned} \tag{3.412}$$

or

$$\dot{p}_1 = qE_1 + \frac{q}{c}\left(\dot{\mathbf{r}} \times \mathbf{B} \right)_1 . \tag{3.413}$$

Similar equations are obtained for $\dot{p}_2$ and $\dot{p}_3$ in the same way. These 3 equations can be written in a 3-vector form as

$$\dot{\mathbf{p}} = q\left(\mathbf{E} + \frac{1}{c}\dot{\mathbf{r}} \times \mathbf{B} \right) , \tag{3.414}$$

where

$$\mathbf{p} = m\gamma\dot{\mathbf{r}} . \tag{3.415}$$

Alternatively, Eq. (3.414) can be rewritten as [see Eq. (3.415)]

$$\ddot{\mathbf{r}} = \frac{q}{m\gamma}\left(\mathbf{E} + \frac{1}{c}\dot{\mathbf{r}} \times \mathbf{B} \right) - \frac{\dot{\gamma}\dot{\mathbf{r}}}{\gamma} , \tag{3.416}$$

where [see Eq. (3.410)]

$$\frac{\dot{\gamma}}{\gamma} = \frac{1}{c}\frac{\frac{u}{c}}{1 - \frac{u^2}{c^2}}\dot{u} , \tag{3.417}$$

and where

$$u = |\dot{\mathbf{r}}| \ . \tag{3.418}$$

With the help of Eqs. (3.414) and (3.415) and the relation

$$\frac{\mathrm{d}\mathbf{p}^2}{\mathrm{d}t} = 2\mathbf{p}\cdot\dot{\mathbf{p}} \ , \tag{3.419}$$

one finds [by multiplying Eq. (3.414) from the left by $\mathbf{p}$] that

$$\frac{1}{2}\frac{\mathrm{d}\mathbf{p}^2}{\mathrm{d}t} = \mathbf{p}\cdot\dot{\mathbf{p}} = qm\gamma\dot{\mathbf{r}}\cdot\mathbf{E} \ , \tag{3.420}$$

thus [see Eq. (3.410)]

$$\dot{u} = \frac{q\left(1-\frac{u^2}{c^2}\right)}{mu\gamma}\dot{\mathbf{r}}\cdot\mathbf{E} \ . \tag{3.421}$$

Combining Eqs. (3.416), (3.417) and (3.421) yields

$$\ddot{\mathbf{r}} = \frac{q}{m\gamma}\left(\mathbf{E}+\frac{1}{c}\dot{\mathbf{r}}\times\mathbf{B}-\frac{(\dot{\mathbf{r}}\cdot\mathbf{E})\,\dot{\mathbf{r}}}{c^2}\right) \ . \tag{3.422}$$

20. Recall that the equation of motion in a 3-vector form is given by Eq. (3.414).
    a) In general, the force 4-vector $F = \mathrm{d}P/\mathrm{d}\tau$ (1.77), where $P = m\mathrm{d}X/\mathrm{d}\tau$ [see Eq. (1.64)], is related to the force 3-vector $\mathbf{f} = \mathrm{d}\mathbf{p}/\mathrm{d}t$ (1.79) by [see Eq. (1.81)]

$$F = \gamma\left(\frac{\mathbf{f}\cdot\dot{\mathbf{r}}}{c},\mathbf{f}\right)^{\mathrm{T}} \ , \tag{3.423}$$

where $\dot{\mathbf{r}} = \mathrm{d}\mathbf{r}/\mathrm{d}t$ 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. (3.414)]

$$\mathbf{f} = q\left(\mathbf{E}+\frac{1}{c}\dot{\mathbf{r}}\times\mathbf{B}\right) \ , \tag{3.424}$$

and thus with the help of the relation (1.68), which is given by

$$\frac{1}{\mathrm{d}\tau} = \gamma\frac{1}{\mathrm{d}t} \ , \tag{3.425}$$

one finds that

$$F = \frac{q}{c}\eta\hat{F}\frac{\mathrm{d}X}{\mathrm{d}\tau} \ , \tag{3.426}$$

where $\hat{F}$ is given by [see Eq. (2.29)]

$$\hat{F} = \begin{pmatrix} 0 & E_1 & E_2 & E_3 \\ -E_1 & 0 & -B_3 & B_2 \\ -E_2 & B_3 & 0 & -B_1 \\ -E_3 & -B_2 & B_1 & 0 \end{pmatrix} , \tag{3.427}$$

and the Minkowski metric η is given by [see Eq. (1.14)]

$$\eta = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} , \tag{3.428}$$

or

$$\frac{\mathrm{d}U}{\mathrm{d}\tau} = \frac{q\eta\hat{F}}{mc} U , \tag{3.429}$$

where

$$U = \frac{\mathrm{d}X}{\mathrm{d}\tau} . \tag{3.430}$$

Note that the right hand side of Eq. (3.429) is transformed according to the Lorentz transformation, i.e. [see Eqs. (1.11) and (2.32)]

$$\Lambda\eta\hat{F}U = \eta\hat{F}'U' . \tag{3.431}$$

b) For the case where $\mathbf{E} = E_1\hat{\mathbf{x}}_1$ and $\mathbf{B} = B_1\hat{\mathbf{x}}_1$ one has

$$\frac{q\eta\hat{F}}{mc} = \frac{q}{mc} \begin{pmatrix} 0 & E_1 & 0 & 0 \\ E_1 & 0 & 0 & 0 \\ 0 & 0 & 0 & B_1 \\ 0 & 0 & -B_1 & 0 \end{pmatrix} , \tag{3.432}$$

and thus [see Eq. (3.429)]

$$\frac{\mathrm{d}}{\mathrm{d}\tau} \begin{pmatrix} U_0 \\ U_1 \end{pmatrix} = \frac{qE_1}{mc} \sigma_{\mathrm{E}} \begin{pmatrix} U_0 \\ U_1 \end{pmatrix} , \tag{3.433}$$

$$\frac{\mathrm{d}}{\mathrm{d}\tau} \begin{pmatrix} U_2 \\ U_3 \end{pmatrix} = \frac{qB_1}{mc} \sigma_{\mathrm{B}} \begin{pmatrix} U_2 \\ U_3 \end{pmatrix} , \tag{3.434}$$

where the 2×2 matrices σ_{E} and σ_{B}, which are given by

$$\sigma_{\mathrm{E}} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} , \tag{3.435}$$

$$\sigma_{\mathrm{B}} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} , \tag{3.436}$$

satisfy the relation $\sigma_{\mathrm{E}}^2 = -\sigma_{\mathrm{B}}^2 = \mathbf{1}$, where $\mathbf{1}$ is the 2×2 identity matrix. The solution is thus given by

$$\begin{pmatrix} U_0(\tau) \\ U_1(\tau) \end{pmatrix} = \exp\left(\frac{qE_1\tau}{mc}\sigma_{\mathrm{E}}\right) \begin{pmatrix} U_0(0) \\ U_1(0) \end{pmatrix} , \tag{3.437}$$

$$\begin{pmatrix} U_2(\tau) \\ U_3(\tau) \end{pmatrix} = \exp\left(\frac{qB_1\tau}{mc}\sigma_{\mathrm{B}}\right) \begin{pmatrix} U_2(0) \\ U_3(0) \end{pmatrix} . \tag{3.438}$$

With the help of the Taylor expansion

$$\exp(M) = 1 + M + \frac{M^2}{2!} + \frac{M^3}{3!} + \frac{M^4}{4!} + \cdots \ , \tag{3.439}$$

one obtains

$$\begin{pmatrix} U_0(\tau) \\ U_1(\tau) \end{pmatrix} = \begin{pmatrix} \cosh \frac{qE_1\tau}{mc} & \sinh \frac{qE_1\tau}{mc} \\ \sinh \frac{qE_1\tau}{mc} & \cosh \frac{qE_1\tau}{mc} \end{pmatrix} \begin{pmatrix} U_0(0) \\ U_1(0) \end{pmatrix} \ , \tag{3.440}$$

$$\begin{pmatrix} U_2(\tau) \\ U_3(\tau) \end{pmatrix} = \begin{pmatrix} \cos \frac{qB_1\tau}{mc} & \sin \frac{qB_1\tau}{mc} \\ -\sin \frac{qB_1\tau}{mc} & \cos \frac{qB_1\tau}{mc} \end{pmatrix} \begin{pmatrix} U_2(0) \\ U_3(0) \end{pmatrix} \ . \tag{3.441}$$

As can be seen from the comparison with Eq. (1.201), 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, ϕ, z). For simplicity, it is assumed that the angle ϕ 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

$$\frac{\mathrm{d}\mathbf{r}}{\mathrm{d}s} = (\sin(\theta(s)), 0, \cos(\theta(s))) \ . \tag{4.1}$$

In the paraxial approximation it is assumed that the angle θ is small.

4.1.1 ABCD Matrix

Let $r(z)$ be the radial coordinate of an optical ray and let $r' = \mathrm{d}r/\mathrm{d}z$. When the transformation from the input values $r_{\mathrm{in}} = r(z_{\mathrm{in}})$ and $r'_{\mathrm{in}} = r'(z_{\mathrm{in}})$ to the output values $r_{\mathrm{out}} = r(z_{\mathrm{out}})$ and $r'_{\mathrm{out}} = r'(z_{\mathrm{out}})$ is found to be a linear one, it can be expressed by the so-called ABCD ray matrix

$$\begin{pmatrix} r_{\mathrm{out}} \\ r'_{\mathrm{out}} \end{pmatrix} = \begin{pmatrix} A & B \\ C & D \end{pmatrix} \begin{pmatrix} r_{\mathrm{in}} \\ r'_{\mathrm{in}} \end{pmatrix} \ . \tag{4.2}$$

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)]

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix} = \begin{pmatrix} 1 & d \\ 0 & 1 \end{pmatrix} , \tag{4.3}$$

where $d = z_{\mathrm{out}} - z_{\mathrm{in}}$. (b) In the paraxial approximation the incidence and transmission angles θ_{i} and θ_{t} are both assumed small, and consequently the Snell's law (3.55), which is given by

$$n_1 \sin\theta_{\mathrm{i}} = n_2 \sin\theta_{\mathrm{t}} , \tag{4.4}$$

can be approximated by the relation

$$n_1\theta_{\mathrm{i}} = n_2\theta_{\mathrm{t}} . \tag{4.5}$$

In this approximation the ABCD matrix is given by [see Fig. 4.1(b)]

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & \frac{n_1}{n_2} \end{pmatrix} . \tag{4.6}$$

(c) For this case [see Fig. 4.1(c)]

$$\theta_{\mathrm{i}} \simeq r'_{\mathrm{in}} + \frac{r_{\mathrm{in}}}{R} , \tag{4.7}$$

$$\theta_{\mathrm{t}} \simeq r'_{\mathrm{out}} + \frac{r_{\mathrm{in}}}{R} , \tag{4.8}$$

and thus [see Eqs. (4.5)]

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ \frac{n_1 - n_2}{n_2 R} & \frac{n_1}{n_2} \end{pmatrix} . \tag{4.9}$$

(d) For a lens of focal length f the focusing condition implies that

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ -\frac{1}{f} & 1 \end{pmatrix} . \tag{4.10}$$

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)]

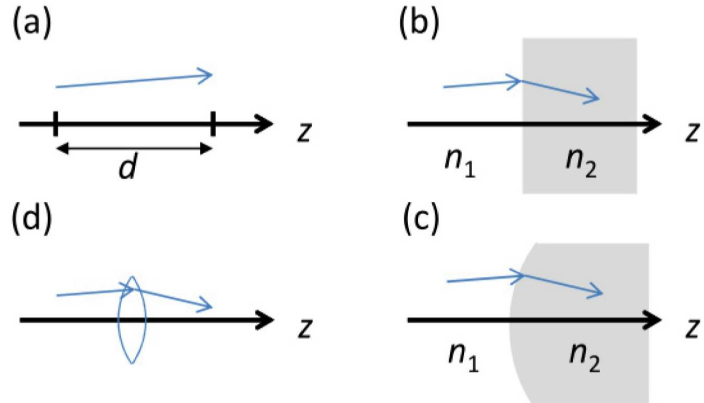

Fig. 4.1. The ABCD ray matrix.

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ -\frac{n-1}{R_2} & n \end{pmatrix} \begin{pmatrix} 1 & 0 \\ \frac{1-n}{nR_1} & \frac{1}{n} \end{pmatrix} \tag{4.11}$$

$$= \begin{pmatrix} 1 & 0 \\ -\frac{1}{f} & 1 \end{pmatrix} , \tag{4.12}$$

$$\tag{4.13}$$

where the focal length f is given by the so-called lensmaker's equation

$$\frac{1}{f} = (n-1)\left(\frac{1}{R_1} - \frac{1}{R_2}\right) . \tag{4.14}$$

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 $A = D$. The underlying symmetry property that is responsible for these properties is discussed below.

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)]

$$\begin{pmatrix} r_{\text{out}} \\ r'_{\text{out}} \end{pmatrix} = \begin{pmatrix} A & B \\ C & D \end{pmatrix} \begin{pmatrix} r_{\text{in}} \\ r'_{\text{in}} \end{pmatrix} . \tag{4.15}$$

The inversion symmetry that the optical element is assumed to posses implies that the following must hold (note that when reversing the direction of an optical ray $r \rightarrow r$ and $r' \rightarrow -r'$)

$$\begin{pmatrix} r_{\mathrm{in}} \\ -r'_{\mathrm{in}} \end{pmatrix} = \begin{pmatrix} A & B \\ C & D \end{pmatrix} \begin{pmatrix} r_{\mathrm{out}} \\ -r'_{\mathrm{out}} \end{pmatrix} , \tag{4.16}$$

thus

$$\begin{pmatrix} r_{\mathrm{out}} \\ r'_{\mathrm{out}} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}^{-1} \begin{pmatrix} A & B \\ C & D \end{pmatrix}^{-1} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} r_{\mathrm{in}} \\ r'_{\mathrm{in}} \end{pmatrix} . \tag{4.17}$$

The requirement that both Eqs. (4.15) and (4.17) must hold for any $(r_{\mathrm{in}}, r'_{\mathrm{in}})^T$ implies that

$$\begin{aligned} \begin{pmatrix} A & B \\ C & D \end{pmatrix} &= \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}^{-1} \begin{pmatrix} A & B \\ C & D \end{pmatrix}^{-1} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \\ &= \frac{1}{AD - BC} \begin{pmatrix} D & B \\ C & A \end{pmatrix} , \end{aligned} \tag{4.18}$$

thus (unless $B = C = 0$)

$$AD - BC = \det M = 1 , \tag{4.19}$$

and

$$A = D . \tag{4.20}$$

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

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix} = M_0^N , \tag{4.21}$$

where

$$M_0 = M_2 M_1 \ , \tag{4.22}$$

and where M_1 and M_2 are given by [see Eqs. (4.3) and (4.10)]

$$M_n = \begin{pmatrix} 1 & 0 \\ -\frac{2}{R_n} & 1 \end{pmatrix} \begin{pmatrix} 1 & d \\ 0 & 1 \end{pmatrix} \ . \tag{4.23}$$

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

$$\lambda_\pm = \frac{\mathrm{Tr}(M_0)}{2} \pm \sqrt{\left(\frac{\mathrm{Tr}(M_0)}{2}\right)^2 - 1} \ . \tag{4.24}$$

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

$$\left|\frac{\mathrm{Tr}(M_0)}{2}\right| \leq 1 \ . \tag{4.25}$$

The trace of M_0 is given by

$$\mathrm{Tr}(M_0) = 2 - \frac{4d}{R_1} - \frac{4d}{R_2} + \frac{4d^2}{R_1 R_2} \ , \tag{4.26}$$

and thus the condition (4.25) can be expressed as

$$0 \leq \frac{\mathrm{Tr}(M_0) + 2}{4} = g_1 g_2 \leq 1 \ , \tag{4.27}$$

where

$$g_n = 1 - \frac{d}{R_n} \ . \tag{4.28}$$

4.1.2 Möbius Transformation

The intersection points z_{in} and z_{out} of the (extrapolated) input and output rays, respectively, with the optical axis are given by

$$z_{\mathrm{in}} = \frac{r_{\mathrm{in}}}{r'_{\mathrm{in}}} \ , \tag{4.29}$$

$$z_{\mathrm{out}} = \frac{r_{\mathrm{out}}}{r'_{\mathrm{out}}} \ . \tag{4.30}$$

The following holds

$$z_{\mathrm{out}} = \frac{r_{\mathrm{out}}}{r'_{\mathrm{out}}} = \frac{A r_{\mathrm{in}} + B r'_{\mathrm{in}}}{C r_{\mathrm{in}} + D r'_{\mathrm{in}}} = f_{A,B,C,D}(z_{\mathrm{in}}) \ , \tag{4.31}$$

where

$$f_{A,B,C,D}\left(z_{\text{in}}\right)=\frac{Az_{\text{in}}+B}{Cz_{\text{in}}+D}\ . \tag{4.32}$$

Note that $f_{A,B,C,D}\left(z_{\text{in}}\right)$ is a Möbius transformation.

Claim. The following holds

$$f_{A_2,B_2,C_2,D_2}\left(f_{A_1,B_1,C_1,D_1}\left(z\right)\right)=f_{A,B,C,D}\left(z\right)\ , \tag{4.33}$$

where

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix}=\begin{pmatrix} A_2 & B_2 \\ C_2 & D_2 \end{pmatrix}\begin{pmatrix} A_1 & B_1 \\ C_1 & D_1 \end{pmatrix}\ . \tag{4.34}$$

Proof. With the help of Eq. (4.32) one finds that

$$\begin{aligned} f_{A_2,B_2,C_2,D_2}\left(f_{A_1,B_1,C_1,D_1}\left(z\right)\right) &= \frac{A_2\frac{A_1z+B_1}{C_1z+D_1}+B_2}{C_2\frac{A_1z+B_1}{C_1z+D_1}+D_2} \\ &= \frac{\left(A_2A_1+B_2C_1\right)z+A_2B_1+B_2D_1}{\left(C_2A_1+D_2C_1\right)z+C_2B_1+D_2D_1}\ , \end{aligned} \tag{4.35}$$

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. (3.111), (3.112) and (3.113).

Exercise 4.1.5. Consider a planar optical ray, for which ϕ is assumed to be a constant. Show that

$$\frac{\mathrm{d}^2r}{\mathrm{d}z^2}=\frac{1}{2n_{\mathrm{c}}^2}\frac{\mathrm{d}n^2}{\mathrm{d}r}\ , \tag{4.36}$$

where n_{c} is a constant.

Solution 4.1.5. With the help of Eq. (3.113)

$$0=\frac{\mathrm{d}}{\mathrm{d}s}\left(n\frac{\mathrm{d}z}{\mathrm{d}s}\right)\ , \tag{4.37}$$

one finds that

$$n\frac{\mathrm{d}z}{\mathrm{d}s}=n_{\mathrm{c}}\ , \tag{4.38}$$

where n_{c} is a constant. By employing the relation [see Eq. (4.38)]

$$\frac{\mathrm{d}}{\mathrm{d}s} = \frac{\mathrm{d}z}{\mathrm{d}s}\frac{\mathrm{d}}{\mathrm{d}z} = \frac{n_{\mathrm{c}}}{n}\frac{\mathrm{d}}{\mathrm{d}z} , \tag{4.39}$$

Eq. (3.111), which for the case of a constant ϕ is given by

$$\frac{\mathrm{d}n}{\mathrm{d}r} = \frac{\mathrm{d}}{\mathrm{d}s}\left(n\frac{\mathrm{d}r}{\mathrm{d}s}\right) , \tag{4.40}$$

becomes

$$\frac{\mathrm{d}n}{\mathrm{d}r} = \frac{n_{\mathrm{c}}^2}{n}\frac{\mathrm{d}^2 r}{\mathrm{d}z^2} , \tag{4.41}$$

or

$$\frac{1}{2n_{\mathrm{c}}^2}\frac{\mathrm{d}n^2}{\mathrm{d}r} = \frac{\mathrm{d}^2 r}{\mathrm{d}z^2} . \tag{4.42}$$

Exercise 4.1.6. Consider a planar optical ray, for which ϕ is assumed to be a constant. Show that

$$\left(\frac{\mathrm{d}r}{\mathrm{d}z}\right)^2 = \frac{n^2 - n_{\mathrm{c}}^2}{n_{\mathrm{c}}^2} , \tag{4.43}$$

where n_{c} is a constant.

Solution 4.1.6. When ϕ is a constant the requirement that $|\mathrm{d}\mathbf{r}/\mathrm{d}s| = 1$ implies that [see Eq. (3.100)]

$$1 = \left(\frac{\mathrm{d}r}{\mathrm{d}s}\right)^2 + \left(\frac{\mathrm{d}z}{\mathrm{d}s}\right)^2 . \tag{4.44}$$

Multiplying Eq. (4.36) by $\mathrm{d}r/\mathrm{d}z$ yields

$$\frac{1}{2n_{\mathrm{c}}^2}\frac{\mathrm{d}n^2}{\mathrm{d}z} = \frac{\mathrm{d}r}{\mathrm{d}z}\frac{\mathrm{d}^2 r}{\mathrm{d}z^2} , \tag{4.45}$$

or

$$\frac{\mathrm{d}}{\mathrm{d}z}\left(\left(\frac{\mathrm{d}r}{\mathrm{d}z}\right)^2 - \frac{n^2}{n_{\mathrm{c}}^2}\right) = 0 , \tag{4.46}$$

thus

$$\left(\frac{\mathrm{d}r}{\mathrm{d}z}\right)^2 - \frac{n^2}{n_{\mathrm{c}}^2} = \mathcal{C} , \tag{4.47}$$

where $\mathcal{C}$ is a constant. On the other hand [see Eqs. (4.38) and (4.44)]

$$\begin{aligned}\left(\frac{\mathrm{d}r}{\mathrm{d}z}\right)^2-\frac{n^2}{n_{\mathrm{c}}^2}&=\left(\frac{\frac{\mathrm{d}r}{\mathrm{d}s}}{\frac{\mathrm{d}z}{\mathrm{d}s}}\right)^2-\frac{n^2}{n_{\mathrm{c}}^2}\\&=\frac{1-\frac{n_{\mathrm{c}}^2}{n^2}}{\frac{n_{\mathrm{c}}^2}{n^2}}-\frac{n^2}{n_{\mathrm{c}}^2}\\&=-1\ ,\end{aligned}\tag{4.48}$$

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

$$n_{\mathrm{GRIN}}\left(r\right)=n_0\sqrt{1-g^2r^2}\ ,\tag{4.49}$$

where both n_0 and g are constants.

Claim. The ABCD ray matrix of a GRIN medium is given by

$$\begin{pmatrix}A & B\\ C & D\end{pmatrix}=\begin{pmatrix}\cos\left(gz\right) & \frac{1}{g}\sin\left(gz\right)\\ -g\sin\left(gz\right) & \cos\left(gz\right)\end{pmatrix}\ .\tag{4.50}$$

Proof. With the help of Eqs. (4.36) and (4.49) one finds that

$$\frac{\mathrm{d}^2r}{\mathrm{d}z^2}=-\frac{n_0^2}{n_{\mathrm{c}}^2}g^2r\ .\tag{4.51}$$

Recall that the constant n_{c} is given by $n_{\mathrm{c}}=n\cos\theta$ [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

$$\frac{\mathrm{d}^2r}{\mathrm{d}z^2}=-g^2r\ .\tag{4.52}$$

The solution thus reads

$$r\left(z\right)=r_0\cos\left(gz\right)+\frac{r_0'}{g}\sin\left(gz\right)\ ,\tag{4.53}$$

where both r_0 and r_0' are constants. The derivative $r'=\mathrm{d}r/\mathrm{d}z$ is given by

$$r'\left(z\right)=-gr_0\sin\left(gz\right)+r_0'\cos\left(gz\right)\ .\tag{4.54}$$

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

$$p=\frac{2\pi}{g}\ .\tag{4.55}$$

4.2 Paraxial Waves

Recall that in the scalar approximation all three components of the vector fields **E** and **H** satisfy the Helmholtz equation (2.151), which is given by

$$\left(\boldsymbol{\nabla}^2 + n^2 k_0^2\right)\psi = 0 \; , \tag{4.56}$$

where n is the refractive index and where [see Eq. (2.142)]

$$k_0 = \frac{\omega}{c} \; . \tag{4.57}$$

4.2.1 Paraxial Approximation

Consider a solution having the form

$$\psi = A\left(x, y, z\right) e^{i n_{\mathrm{a}} k_0 z} \; , \tag{4.58}$$

where the constant n_{a} represents a characteristic value of the refractive index n in the medium. Substituting into Eq. (4.56) yields

$$\boldsymbol{\nabla}_{\perp}^2 A + \frac{\partial}{\partial z}\left(\frac{\partial A}{\partial z} + 2 i n_{\mathrm{a}} k_0 A\right) + \left(n^2 - n_{\mathrm{a}}^2\right) k_0^2 A = 0 \; , \tag{4.59}$$

where

$$\boldsymbol{\nabla}_{\perp}^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} \; . \tag{4.60}$$

In the paraxial approximation, which assumes

$$\left|\frac{\partial A}{\partial z}\right| \ll 2 n_{\mathrm{a}} k_0 \left|A\right| \; ,$$

this becomes

$$i\frac{\partial A}{\partial z} = \mathcal{H} A \; , \tag{4.61}$$

where

$$\mathcal{H} = -\frac{1}{2 n_{\mathrm{a}} k_0}\boldsymbol{\nabla}_{\perp}^2 - \frac{\left(n^2 - n_{\mathrm{a}}^2\right) k_0}{2 n_{\mathrm{a}}} \; . \tag{4.62}$$

4.2.2 Gaussian Beam in GRIN Medium

Consider a GRIN medium, whose refractive index n_{GRIN} is given by Eq. (4.49). For this case Eq. (4.61) becomes (the constant n_{a} is chosen to be given by $n_{\mathrm{a}} = n_0$)

$$i\frac{\partial A}{\partial z} = \left(-\frac{1}{2n_0k_0}\boldsymbol{\nabla}_\perp^2 + \frac{n_0k_0g^2r^2}{2}\right)A \; . \tag{4.63}$$

Consider a gaussian beam solution, for which A has the form

$$A = A_0 e^{-i\left(P(z)-\frac{n_0k_0r^2}{2q(z)}\right)} \; , \tag{4.64}$$

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

$$q(z) = \frac{Aq_0 + B}{Cq_0 + D} \; , \tag{4.66}$$

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, ϕ, z) the following holds

$$\boldsymbol{\nabla}_\perp^2 = \frac{1}{r}\frac{\partial}{\partial r}r\frac{\partial}{\partial r} + \frac{1}{r^2}\frac{\partial^2}{\partial \phi^2} \; , \tag{4.68}$$

and thus Eq. (4.63) yields

$$\frac{\mathrm{d}P}{\mathrm{d}z} + \frac{i}{q} + \frac{n_0k_0r^2}{2q^2}\left(\frac{\mathrm{d}q}{\mathrm{d}z} - 1 - g^2q^2\right) = 0 \; . \tag{4.69}$$

The above must hold for every r, and thus

$$\frac{\mathrm{d}q}{\mathrm{d}z} = 1 + g^2q^2 \; , \tag{4.70}$$

and

$$\frac{\mathrm{d}P}{\mathrm{d}z} = -\frac{i}{q} \; . \tag{4.71}$$

The general solution of Eq. (4.70) reads

$$q(z) = \frac{\tan(gz + C)}{g} , \tag{4.72}$$

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

$$q(z) = \frac{\cos(gz) q_0 + \frac{1}{g}\sin(gz)}{-g\sin(gz) q_0 + \cos(gz)} , \tag{4.75}$$

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 $1/R(z = 0) = 0$ [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

$$\frac{1}{q(z)} \simeq \frac{1}{R(z)} \simeq \frac{1}{z} . \tag{4.80}$$

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_{out} of an output ray to the the intersection point z_{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

$$q(z) = \frac{Aq_0 + B}{Cq_0 + D} , \tag{4.81}$$

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

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix} = \begin{pmatrix} 1 & d \\ 0 & 1 \end{pmatrix} , \tag{4.82}$$

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

$$n(x, y, z) = \sqrt{n_0^2(x, y) + n_{\mathrm{p}}^2} , \tag{4.83}$$

where the term $n_{\mathrm{p}} = n_{\mathrm{p}}(x, y, z)$ is considered as a perturbation. In the scalar approximation the field components are required to satisfy the Helmholtz equation (4.56)

$$\left(\boldsymbol{\nabla}_{\perp}^2 + \frac{\mathrm{d}^2}{\mathrm{d}z^2} + k_0^2 n^2 \right) \psi = 0 , \tag{4.84}$$

where $\boldsymbol{\nabla}_{\perp}^2$ is given by Eq. (4.60), $k_0 = \omega/c$, ω 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 $n_{\mathrm{p}} = 0$, are given by

$$\psi_{\pm}(x,y,z) = \psi_0(x,y)\, e^{\pm i\beta z} \,. \tag{4.85}$$

The dispersion relation $\beta(\omega)$ in that region is assumed to be linear and approximately given by

$$\beta(\omega) = \frac{\omega n_{\mathrm{eff}}}{c} = k_0 n_{\mathrm{eff}} \,, \tag{4.86}$$

where n_{eff} is the mode effective refractive index (near ω_0). This assumption implies, as can be seen from Eq. (4.84), that the function $\psi_0(x,y)$ satisfies the following transverse equation

$$\left\{\boldsymbol{\nabla}_{\perp}^2 + k_0^2\left[n_0^2(x,y) - n_{\mathrm{eff}}^2\right]\right\}\psi_0 = 0 \,. \tag{4.87}$$

Consider a solution to the perturbed problem having the form

$$\psi(x,y,z) = \psi_0(x,y)\left[A_+(z)\, e^{i\beta z} + A_-(z)\, e^{-i\beta z}\right] \,. \tag{4.88}$$

Substituting into Eq. (4.84) and employing Eq. (4.87) yields

$$\left[\frac{\mathrm{d}^2}{\mathrm{d}z^2} + k_0^2\left(n_{\mathrm{p}}^2 + n_{\mathrm{eff}}^2\right)\right]\psi_0\left(A_+ e^{i\beta z} + A_- e^{-i\beta z}\right) = 0 \,. \tag{4.89}$$

The envelope functions $A_{\pm}(z)$, which become constants in the unperturbed case, are assume to be nearly constant on the length scale of a single wavelength, and therefore

$$\frac{\mathrm{d}^2}{\mathrm{d}z^2}\left(A_{\pm} e^{\pm i\beta z}\right) \simeq \left(\pm 2i\beta\frac{\mathrm{d}A_{\pm}}{\mathrm{d}z} - \beta^2 A_{\pm}\right) e^{\pm i\beta z} \,. \tag{4.90}$$

Employing this approximation in Eq. (4.89) leads to

$$\begin{aligned} &2i\beta\left(\frac{\mathrm{d}A_+}{\mathrm{d}z} e^{i\beta z} - \frac{\mathrm{d}A_-}{\mathrm{d}z} e^{-i\beta z}\right)\psi_0 \\ &+ \left[k_0^2\left(n_{\mathrm{p}}^2 + n_{\mathrm{eff}}^2\right) - \beta^2\right]\psi_0\left(A_+ e^{i\beta z} + A_- e^{-i\beta z}\right) = 0 \,. \end{aligned} \tag{4.91}$$

Multiplying by ψ_0^* and integrating over the xy plane yield

$$\begin{aligned} &2i\beta\left(\frac{\mathrm{d}A_+}{\mathrm{d}z} e^{i\beta z} - \frac{\mathrm{d}A_-}{\mathrm{d}z} e^{-i\beta z}\right) \\ &+ \left[k_0^2 n_{\mathrm{eff}}^2(1+2D) - \beta^2\right]\left(A_+ e^{i\beta z} + A_- e^{-i\beta z}\right) = 0 \,, \end{aligned} \tag{4.92}$$

where

$$D(z) = \frac{\frac{1}{2n_{\mathrm{eff}}^2}\int\int \mathrm{d}x\mathrm{d}y\ \left|\psi_0\right|^2 n_{\mathrm{p}}^2}{\int\int \mathrm{d}x\mathrm{d}y\ \left|\psi_0\right|^2} \,. \tag{4.93}$$

The coupling term $D(z)$ for the case of Bragg grating is assumed to have the form

$$D(z) = \mathcal{D}(z)\, e^{\frac{2\pi i z}{\Lambda}} + \mathcal{D}^*(z)\, e^{-\frac{2\pi i z}{\Lambda}} \ , \tag{4.94}$$

where the envelope function $\mathcal{D}(z)$ is assumed to be nearly constant on the length scale of the grating period Λ. The optical angular frequency ω is assumed to be close to the Bragg frequency ω_{B} (i.e. ω_0 is taken to be equal to ω_{B}), which is given by

$$\omega_{\mathrm{B}} = \frac{\pi c}{\Lambda n_{\mathrm{eff}}} = c\frac{2\pi}{\lambda_{\mathrm{B}}} \ , \tag{4.95}$$

were $\lambda_{\mathrm{B}} = 2n_{\mathrm{eff}}\Lambda$ is the Bragg wavelength, i.e. $|\omega - \omega_{\mathrm{B}}| \ll \omega_{\mathrm{B}}$. Thus the factor β can be expressed in terms of the normalized detuning factor δ as

$$\beta = \frac{\pi}{\Lambda}(1-\delta) \ , \tag{4.96}$$

where

$$\delta = -\frac{\omega - \omega_{\mathrm{B}}}{\omega_{\mathrm{B}}} \simeq \frac{\Delta_\lambda}{\lambda_{\mathrm{B}}} \ll 1 \ , \tag{4.97}$$

and where $\Delta_\lambda = \lambda - \lambda_{\mathrm{B}}$ is the offset wavelength. To first order in δ

$$k_0^2 n_{\mathrm{eff}}^2 - \beta^2 \simeq 2\left(\frac{\pi}{\Lambda}\right)^2 \delta \ . \tag{4.98}$$

With these notations and approximations one obtains

$$\begin{aligned}&\left(\frac{\mathrm{d}A_+}{\mathrm{d}z} - \frac{\pi i \delta}{\Lambda}A_+\right)e^{i\beta z} - \left(\frac{\mathrm{d}A_-}{\mathrm{d}z} + \frac{\pi i \delta}{\Lambda}A_-\right)e^{-i\beta z} \\ &-\frac{\pi i}{\Lambda}D\left(A_+e^{i\beta z} + A_-e^{-i\beta z}\right) = 0 \ .\end{aligned} \tag{4.99}$$

Collecting terms oscillating close to $\exp(\pi i z/\Lambda)$ yields

$$\frac{\mathrm{d}A_+}{\mathrm{d}z} - \frac{\pi i \delta}{\Lambda}A_+ - \frac{\pi i}{\Lambda}\mathcal{D}e^{\frac{2\pi i \delta z}{\Lambda}}A_- = 0 \ , \tag{4.100}$$

whereas collecting terms oscillating close to $\exp(-\pi i z/\Lambda)$ yields

$$\frac{\mathrm{d}A_-}{\mathrm{d}z} + \frac{\pi i \delta}{\Lambda}A_- + \frac{\pi i}{\Lambda}\mathcal{D}^*e^{-\frac{2\pi i \delta z}{\Lambda}}A_+ = 0 \ . \tag{4.101}$$

In order to ensure that small terms are kept only up to first order (recall that both $\mathcal{D}$ and δ are considered to be small) the terms $\exp(\pm 2\pi i \delta z/\Lambda)$ are replaced by unity. In terms of the dimensionless displacement $\zeta = \pi z/\Lambda$ these two coupled equations can be expressed as

$$\frac{dA_+}{d\zeta} - i\delta A_+ - i\mathcal{D}A_- = 0 \; , \tag{4.102}$$

$$\frac{dA_-}{d\zeta} + i\delta A_- + i\mathcal{D}^* A_+ = 0 \; , \tag{4.103}$$

or in matrix form as

$$\frac{d}{d\zeta}\begin{pmatrix} A_+ \\ A_- \end{pmatrix} = \mathcal{M}\begin{pmatrix} A_+ \\ A_- \end{pmatrix} \; , \tag{4.104}$$

where

$$\mathcal{M} = \begin{pmatrix} i\delta & i\mathcal{D} \\ -i\mathcal{D}^* & -i\delta \end{pmatrix} \; . \tag{4.105}$$

In what follows $\mathcal{D}$ is assumed to be a real constant. For this case the solution is given by

$$\begin{pmatrix} A_+(\zeta) \\ A_-(\zeta) \end{pmatrix} = \exp(\mathcal{M}\zeta)\begin{pmatrix} A_+(0) \\ A_-(0) \end{pmatrix} \; . \tag{4.106}$$

For a complex number ζ and for a unit vector $\hat{\mathbf{n}}$ (i.e., $\hat{\mathbf{n}} \cdot \hat{\mathbf{n}} = 1$), the following holds

$$\exp(\zeta \boldsymbol{\sigma} \cdot \hat{\mathbf{n}}) = \cosh\zeta + \boldsymbol{\sigma} \cdot \hat{\mathbf{n}} \sinh\zeta \; , \tag{4.107}$$

where $\boldsymbol{\sigma} = (\sigma_x, \sigma_y, \sigma_z)$ is the vector of Pauli matrices, which are given by

$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \; \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \; \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \; . \tag{4.108}$$

Thus by expressing $\mathcal{M}\zeta$ as

$$\mathcal{M}\zeta = \gamma\zeta(\sigma_x, \sigma_y, \sigma_z) \cdot \begin{pmatrix} 0 \\ \frac{-\mathcal{D}}{\gamma} \\ \frac{i\delta}{\gamma} \end{pmatrix} \; . \tag{4.109}$$

where

$$\gamma = \sqrt{\mathcal{D}^2 - \delta^2} \; , \tag{4.110}$$

one finds that

$$\exp(\mathcal{M}\zeta) = \begin{pmatrix} \cosh(\gamma\zeta) + \frac{i\delta\sinh(\gamma\zeta)}{\gamma} & \frac{i\mathcal{D}\sinh(\gamma\zeta)}{\gamma} \\ -\frac{i\mathcal{D}\sinh(\gamma\zeta)}{\gamma} & \cosh(\gamma\zeta) - \frac{i\delta\sinh(\gamma\zeta)}{\gamma} \end{pmatrix} \; . \tag{4.111}$$

Consider a Bragg grating having length L_{B}. Introducing the grating coupling strength parameter

$$V = \pi N_{\mathrm{B}} \mathcal{D} \,, \tag{4.112}$$

where

$$N_{\mathrm{B}} = \frac{L_{\mathrm{B}}}{\Lambda} \tag{4.113}$$

is the number of periods, and the total detuning factor

$$\Delta = \pi N_{\mathrm{B}} \delta \,, \tag{4.114}$$

where $\delta = \Delta_\lambda / \lambda_{\mathrm{B}}$, the transfer matrix $M_{\mathrm{B}} = \exp(\mathcal{M}\zeta)$ (4.111) can be expresses in terms of the the transmission t_{B} and the reflection r_{B} amplitudes of the FBG as

$$M_{\mathrm{B}} = \begin{pmatrix} \frac{1}{t_{\mathrm{B}}^*} & \frac{r_{\mathrm{B}}}{t_{\mathrm{B}}} \\ \left(\frac{r_{\mathrm{B}}}{t_{\mathrm{B}}}\right)^* & \frac{1}{t_{\mathrm{B}}} \end{pmatrix} \,, \tag{4.115}$$

where

$$t_{\mathrm{B}} = \frac{1}{\cosh\sqrt{V^2 - \Delta^2} - \frac{i\Delta \sinh\sqrt{V^2-\Delta^2}}{\sqrt{V^2-\Delta^2}}} \,, \tag{4.116}$$

$$r_{\mathrm{B}} = \frac{\frac{iV \sinh\sqrt{V^2-\Delta^2}}{\sqrt{V^2-\Delta^2}}}{\cosh\sqrt{V^2 - \Delta^2} - \frac{i\Delta \sinh\sqrt{V^2-\Delta^2}}{\sqrt{V^2-\Delta^2}}} \,. \tag{4.117}$$

The reflection R_{B} and transmission T_{B} probabilities are given by (see Fig. 4.2)

$$R_{\mathrm{B}} = |r_{\mathrm{B}}|^2 = \frac{\frac{V^2 \sinh^2\sqrt{V^2-\Delta^2}}{V^2-\Delta^2}}{1 + \frac{V^2 \sinh^2\sqrt{V^2-\Delta^2}}{V^2-\Delta^2}} \,, \tag{4.118}$$

$$T_{\mathrm{B}} = |t_{\mathrm{B}}|^2 = \frac{1}{1 + \frac{V^2 \sinh^2\sqrt{V^2-\Delta^2}}{V^2-\Delta^2}} \,. \tag{4.119}$$

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 $M = y_2/y_1$.

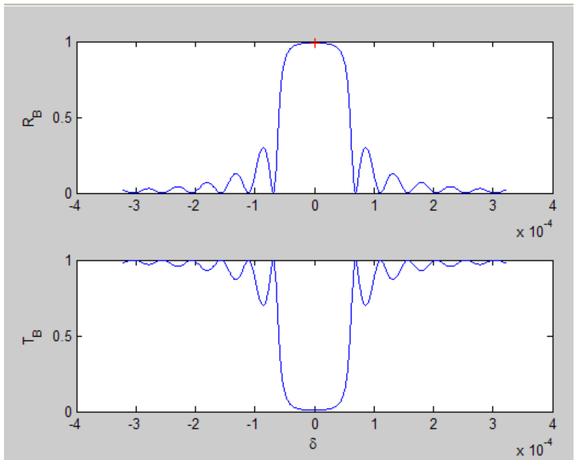

Fig. 4.2. Reflection $R_{\mathrm{B}} = |r_{\mathrm{B}}|^2$ and transmission $T_{\mathrm{B}} = |t_{\mathrm{B}}|^2 = 1 - R_{\mathrm{B}}$ probabilities of a Bragg grating as a function of δ for coupling constant $V = 3$ and $N_{\mathrm{B}} = L_{\mathrm{B}}/\Lambda = 20000$.

2. An optical imaging system is constructed using two thin lenses having focal lengths f_1 and f_2, respectively. The two lenses are attached to each other (with a vanishing gap). The object plane is on the left at a distance s_1 from the lenses, and the image plane is on the right at a distance s_2 from the lenses. The total distance between object plane and image plane is given $s_1+s_2 = s$. Calculate the magnification M of the imaging system.
3. Consider an incident laser spot having a radius R_{in} and a small divergence angle θ_{in}. Employ paraxial ray optics to estimate the output radius R_{out} and output divergence angle θ_{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.
4. **Ball lens** - Find the focal distance of a ball lens having radius R and index of refraction n_{b}.
5. 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.
6. Consider the task of coupling a collimated laser beam of wavelength λ having characteristic mode radius w_{L} into a single mode optical fiber

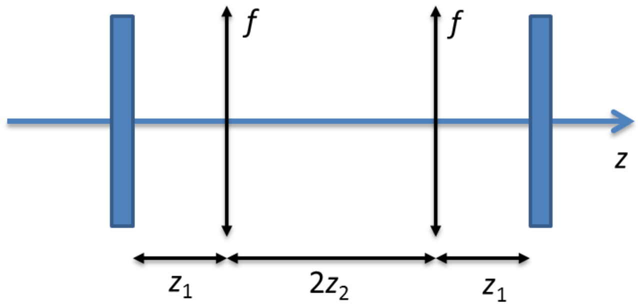

Fig. 4.3. Optical cavity between two flat mirrors with two internal lenses.

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?

7. Consider the cavity seen in Fig. 4.3, which is made of two flat mirrors and two lenses both having focal distance f. The distance between the left (right) mirror and the 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?
8. 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.
9. A gaussian beam illuminates a thin lens having focal length f. Find a relation between the radius of curvature R_1 at the input of the lens, and the output radius of curvature R_2.
10. Find the radius of curvature R corresponding to a complex beam parameter q satisfying the relation

$$q = \frac{Aq + B}{Cq + D} , \tag{4.120}$$

where A, B, C and D are all real.

11. Consider a gaussian beam in free space having wavelength λ. 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_{f} of the new waist that is generated by the focusing effect of the lens and the distance d_{f} between it and the lens.
12. A Gaussian beam having angular frequency ω propagates in vacuum along the z axis. At some point along the axis the spot size is w_1 and the radius of curvature is R_1. Express the spot size at the waist w_0 (i.e. the minimum value of w) as a function of w_1 and R_1.

13. Calculate the modes' frequencies of a single mode fiber ring having length L, with an integrated lump FBG.
14. **Photonic band gap** - Calculate the effective refractive index n_{B} for longitudinal propagation along a Bragg grating.
15. The reflection amplitude r_{B} of a FBG is expressed by Eq. Eq. (4.117) in terms of the grating coupling strength parameter V. Calculate r_{B} in the limit $V \rightarrow \infty$ (assume that the ratio Δ/V remains finite).
16. Evaluate the fixed points of the Möbius transformation (4.32) and determine their stability for the case where the ABCD matrix M can be expressed as

$$M = \begin{pmatrix} 1+\alpha\beta & \beta \\ \alpha & 1 \end{pmatrix} , \tag{4.121}$$

where α and β are complex constants.
17. **Gaussian pulse** - Consider a Gaussian optical pulse having time-dependent amplitude $E(t)$ given by

$$E(t) = E_0 \exp\left(-\gamma t^2 + i\omega_{\mathrm{p}} t\right) , \tag{4.122}$$

where E_0 is a complex constant, $\gamma = \gamma' + i\gamma''$, $\gamma' = \operatorname{Re}\gamma > 0$ determines the width of the pulse, $\gamma'' = \operatorname{Im}\gamma$ represents a linear chirp and the real ω_{p} is the optical angular frequency. Consider two types of transformation. For the first type, which henceforth is referred to as frequency-like, the Fourier amplitude $E(\omega)$, which is related to $E(t)$ by

$$E(\omega) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \mathrm{d}t\; E(t)\, e^{-i\omega t} , \tag{4.123}$$

is transformed according to

$$E(\omega) \rightarrow E'(\omega) = \exp\left(-\frac{(\omega-\omega_{\mathrm{p}})^2}{4\gamma_{\mathrm{F}}}\right) E(\omega) , \tag{4.124}$$

where γ_{F} is a complex constant. For the second type, which is referred to as time-like, the time-domain pulse shape is transformed according to

$$E(t) \rightarrow E'(t) = \exp\left(-\gamma_{\mathrm{T}} t^2\right) E(t) , \tag{4.125}$$

where γ_{T} is a complex constant.
   a) Show that both types can be described in terms of a Möbius transformation mapping the Gaussian pulse variable γ.
   b) Consider a Gaussian optical pulse circulating inside an optical system. In each cycle the pulse first undergoes a time-like transformation with parameter γ_{T}, and then a frequency-like transformation with parameter γ_{F}. Determine the value of the parameter γ is steady state (i.e. after a large number of cycles).

c) **Mode locking** - Assume that the optical system is a cavity formed along the optical axis z between two mirrors. The mirror on the left at $z_1 = -L$ is assumed to be stationary, whereas the mirror on the right at $z_2 = l_{\mathrm{m}} \sin(\omega_{\mathrm{m}} t)$ is assumed to periodically oscillate with a positive amplitude l_{m} and a positive angular frequency ω_{m}. In addition, the cavity contains a section, which provides optical gain. Assume that the effect of this section on a Gaussian optical pulse passing through it can be described in terms of a frequency-like transformation with a fixed positive variable γ_{F} [see Eq. (4.124)]. Calculate the pulse parameter γ in steady state in the limit of slowly moving mirror, for which it is assumed that $\omega_{\mathrm{m}}^2 \ll \gamma_{\mathrm{F}}$.

4.5 Solutions

1. The ABCD matrix is given by [see Eqs. (4.2), (4.3) and (4.10)]

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix} = \begin{pmatrix} 1 & s_2 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ -\frac{1}{f} & 1 \end{pmatrix} \begin{pmatrix} 1 & s_1 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} \frac{f-s_2}{f} & \frac{s_1 f - s_1 s_2 + s_2 f}{f} \\ -\frac{1}{f} & \frac{-s_1+f}{f} \end{pmatrix} . \tag{4.126}$$

Imaging occurs when $B = 0$ (explain why), i.e. when

$$\frac{1}{s_1} + \frac{1}{s_2} = \frac{1}{f} . \tag{4.127}$$

When $B = 0$ the matrix becomes

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix} = \begin{pmatrix} M & 0 \\ -\frac{1}{f} & \frac{1}{M} \end{pmatrix} , \tag{4.128}$$

where

$$M = -\frac{s_2}{s_1} \tag{4.129}$$

is the magnification.

2. With the help of Eq. (4.10) one finds that

$$\begin{pmatrix} 1 & 0 \\ -\frac{1}{f_1} & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ -\frac{1}{f_2} & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ -\frac{1}{f_{\mathrm{e}}} & 1 \end{pmatrix} , \tag{4.130}$$

where f_{e}, which is given by

$$f_{\mathrm{e}} = \frac{f_1 f_2}{f_2 + f_1} , \tag{4.131}$$

is the effective focal length of the two attached lenses. The imaging condition (4.127), which reads

$$\frac{1}{s_1}+\frac{1}{s_2}=\frac{1}{f_{\mathrm{e}}}\;, \tag{4.132}$$

together with the relation $s_1+s_2=s$ lead to

$$\frac{s_1}{s}=\frac{1+\sqrt{1-\frac{4f_{\mathrm{e}}}{s}}}{2}\;, \tag{4.133}$$

$$\frac{s_2}{s}=\frac{1-\sqrt{1-\frac{4f_{\mathrm{e}}}{s}}}{2}\;, \tag{4.134}$$

and thus the magnification is given by [see Eq. (4.129)]

$$M=-\frac{s_2}{s_1}=-\frac{1-\sqrt{1-\frac{4f_{\mathrm{e}}}{s}}}{1+\sqrt{1-\frac{4f_{\mathrm{e}}}{s}}}\;. \tag{4.135}$$

3. With the help of Eqs. (4.2), (4.3) and (4.10) one finds that:
   a) For the case of focusing R_{out} and θ_{out} are given by

$$\begin{aligned}\begin{pmatrix}R_{\mathrm{out}}\\ \theta_{\mathrm{out}}\end{pmatrix}&=\begin{pmatrix}1 & f\\ 0 & 1\end{pmatrix}\begin{pmatrix}1 & 0\\ -\frac{1}{f} & 1\end{pmatrix}\begin{pmatrix}R_{\mathrm{in}}\\ \theta_{\mathrm{in}}\end{pmatrix}\\ &=\begin{pmatrix}f\theta_{\mathrm{in}}\\ \frac{-R_{\mathrm{in}}+f\theta_{\mathrm{in}}}{f}\end{pmatrix}\;,\end{aligned} \tag{4.136}$$

   b) For the case of collimating R_{out} and θ_{out} are given by

$$\begin{aligned}\begin{pmatrix}R_{\mathrm{out}}\\ \theta_{\mathrm{out}}\end{pmatrix}&=\begin{pmatrix}1 & 0\\ -\frac{1}{f} & 1\end{pmatrix}\begin{pmatrix}1 & f\\ 0 & 1\end{pmatrix}\begin{pmatrix}R_{\mathrm{in}}\\ \theta_{\mathrm{in}}\end{pmatrix}\\ &=\begin{pmatrix}R_{\mathrm{in}}+f\theta_{\mathrm{in}}\\ -\frac{R_{\mathrm{in}}}{f}\end{pmatrix}\;,\end{aligned} \tag{4.137}$$

   c) For the case spot expansion the ABCD matrix is given by

$$\begin{aligned}\begin{pmatrix}A & B\\ C & D\end{pmatrix}&=\begin{pmatrix}1 & 0\\ -\frac{1}{f_2} & 1\end{pmatrix}\begin{pmatrix}1 & d\\ 0 & 1\end{pmatrix}\begin{pmatrix}1 & 0\\ -\frac{1}{f_1} & 1\end{pmatrix}\\ &=\begin{pmatrix}-\frac{-f_1+d}{f_1} & d\\ \frac{-f_1+d-f_2}{f_2f_1} & -\frac{d-f_2}{f_2}\end{pmatrix}\;.\end{aligned} \tag{4.138}$$

Collimating occurs when $C = 0$, i.e. when

$$d = f_1 + f_2 \; . \tag{4.139}$$

For this case R_{out} and θ_{out} are given by

$$\begin{pmatrix} R_{\mathrm{out}} \\ \theta_{\mathrm{out}} \end{pmatrix} = \begin{pmatrix} -\frac{f_2}{f_1} & f_1 + f_2 \\ 0 & -\frac{f_1}{f_2} \end{pmatrix} \begin{pmatrix} R_{\mathrm{in}} \\ \theta_{\mathrm{in}} \end{pmatrix} = \begin{pmatrix} -\frac{f_2 R_{\mathrm{in}}}{f_1} + \theta_{\mathrm{in}} d \\ -\frac{f_1}{f_2}\theta_{\mathrm{in}} \end{pmatrix} \; . \tag{4.140}$$

4. For a ball of radius R and index of refraction n_{b} on has [see Eqs. (4.3) and (4.9)]

$$\begin{aligned} &\begin{pmatrix} A & B \\ C & D \end{pmatrix} \\ &= \begin{pmatrix} 1 & 0 \\ \frac{n_{\mathrm{b}}-1}{(-R)} & n_{\mathrm{b}} \end{pmatrix} \begin{pmatrix} 1 & 2R \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ \frac{1-n_{\mathrm{b}}}{n_{\mathrm{b}} R} & \frac{1}{n_{\mathrm{b}}} \end{pmatrix} \\ &= \begin{pmatrix} -1 + \frac{2}{n_{\mathrm{b}}} & \frac{2R}{n_{\mathrm{b}}} \\ \frac{2(1-n_{\mathrm{b}})}{n_{\mathrm{b}} R} & -1 + \frac{2}{n_{\mathrm{b}}} \end{pmatrix} , \end{aligned} \tag{4.141}$$

or

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix} = \begin{pmatrix} 1 & R \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ -\frac{1}{f_{\mathrm{b}}} & 1 \end{pmatrix} \begin{pmatrix} 1 & R \\ 0 & 1 \end{pmatrix} , \tag{4.142}$$

where the focal distance f_{b} is given by

$$f_{\mathrm{b}} = \frac{n_{\mathrm{b}} R}{2\left(n_{\mathrm{b}} - 1\right)} \; . \tag{4.143}$$

5. The ABCD matrix is given by [see Eqs. (4.2), (4.3), (4.6) and (4.50)]

$$\begin{aligned} \begin{pmatrix} A & B \\ C & D \end{pmatrix} &= \begin{pmatrix} 1 & s_1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & n_0 \end{pmatrix} \\ &\times \begin{pmatrix} \cos\theta_{\mathrm{g}} & \frac{1}{g}\sin\theta_{\mathrm{g}} \\ -g\sin\theta_{\mathrm{g}} & \cos\theta_{\mathrm{g}} \end{pmatrix} \\ &\times \begin{pmatrix} 1 & 0 \\ 0 & \frac{1}{n_0} \end{pmatrix} \begin{pmatrix} 1 & s_2 \\ 0 & 1 \end{pmatrix} \\ &= \cos\theta_{\mathrm{g}} \begin{pmatrix} 1 - n_0 g s_1 \tan\theta_{\mathrm{g}} & \frac{n_0 g (s_1+s_2) + \left(1 - n_0^2 g^2 s_1 s_2\right)\tan\theta_{\mathrm{g}}}{g n_0} \\ -n_0 g \tan\theta_{\mathrm{g}} & -n_0 g s_2 \tan\theta_{\mathrm{g}} + 1 \end{pmatrix} . \end{aligned} \tag{4.144}$$

where $\theta_{\mathrm{g}} = gz$ and where [see Eq. (4.55)]

$$g = \frac{2\pi}{p} . \tag{4.145}$$

The displacement s_2 is determined from the imaging condition

$$0 = B = \frac{n_0 g (s_1 + s_2) + \left(1 - n_0^2 g^2 s_1 s_2\right) \tan\theta_{\mathrm{g}}}{g n_0} , \tag{4.146}$$

or

$$0 = \frac{2}{1+T^2} + \frac{1}{q_1} + \frac{1}{q_2} , \tag{4.147}$$

where

$$q_n = n_0 g T s_n - 1 , \tag{4.148}$$

$$T = \tan\frac{\theta_{\mathrm{g}}}{2} , \tag{4.149}$$

$$\frac{2}{1+T^2} = 1 + \cos\theta_{\mathrm{g}} , \tag{4.150}$$

and where $n \in \{1, 2\}$. When this condition is satisfied one has

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix} = \begin{pmatrix} -1 - \frac{2q_1}{1+T^2} & 0 \\ -n_0 g \sin\theta_{\mathrm{g}} & -1 - \frac{2q_2}{1+T^2} \end{pmatrix} , \tag{4.151}$$

or

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix} = \begin{pmatrix} \frac{1}{M} & 0 \\ -\frac{1}{\frac{f_{\mathrm{g}}}{n_0}} & M \end{pmatrix} , \tag{4.152}$$

where the magnification M is given by

$$M = -\frac{q_2}{q_1} . \tag{4.153}$$

6. 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

$$\begin{aligned} \begin{pmatrix} A & B \\ C & D \end{pmatrix} &= \begin{pmatrix} 1 & 0 \\ -\frac{1}{f} & 1 \end{pmatrix} \begin{pmatrix} 1 & f \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & n \end{pmatrix} \\ &= \begin{pmatrix} 1 & fn \\ -\frac{1}{f} & 0 \end{pmatrix} . \end{aligned} \tag{4.154}$$

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)]

$$q_{\mathrm{in}} = -i\frac{\pi w_{\mathrm{F}}^2 n}{\lambda} , \tag{4.155}$$

and the output complex beam parameter q_{out} is given by

$$q_{\mathrm{out}} = -i\frac{\pi w_{\mathrm{L}}^2}{\lambda} . \tag{4.156}$$

Using Eq. (4.66) one finds that

$$q_{\mathrm{out}} = \frac{Aq_{\mathrm{i}} + B}{Cq_{\mathrm{i}} + D} , \tag{4.157}$$

thus

$$i\frac{\pi w_{\mathrm{L}}^2}{\lambda} = f\left(\frac{if\lambda}{\pi w_{\mathrm{F}}^2} + 1\right) . \tag{4.158}$$

The assumption that $w_{\mathrm{F}} \ll \sqrt{f\lambda}$ implies that the real part of the above equation is much smaller than the imaginary part. By neglecting the real part one finds that

$$f = \frac{\pi w_{\mathrm{L}} w_{\mathrm{F}}}{\lambda} , \tag{4.159}$$

or in terms of the diameters $2w_{\mathrm{L}}$ and $2w_{\mathrm{F}}$

$$f = \frac{\pi (2w_{\mathrm{L}}) (2w_{\mathrm{F}})}{4\lambda} . \tag{4.160}$$

7. 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)]

$$M = \begin{pmatrix} 1 & z_1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ -\frac{1}{f} & 1 \end{pmatrix} \begin{pmatrix} 1 & 2z_2 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ -\frac{1}{f} & 1 \end{pmatrix} \begin{pmatrix} 1 & z_1 \\ 0 & 1 \end{pmatrix} . \tag{4.161}$$

The stability condition reads [see Eq. (4.25) and note that $\det M = 1$]

$$\left|\frac{\mathrm{Tr}\,(M)}{2}\right| = \left|1 + \frac{2z_1 z_2}{f}\left(\frac{1}{f} - \frac{z_1 + z_2}{z_1 z_2}\right)\right| \leq 1 , \tag{4.162}$$

and thus

$$f \geq \frac{z_1 z_2}{z_1 + z_2} , \tag{4.163}$$

or

$$\frac{1}{f} \leq \frac{1}{z_1} + \frac{1}{z_2} . \tag{4.164}$$

8. Let z_1 ($z_2 = z_1 + d$) 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 $f = R/2$)

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ -\frac{2}{R} & 1 \end{pmatrix} , \tag{4.165}$$

and thus the complex beam parameter q of a back-reflected gaussian beam q_{out} is related to the parameter q of an incident beam q_{in} by [see Eq. (4.66)]

$$\frac{1}{q_{\mathrm{out}}} = \frac{1}{q_{\mathrm{in}}} - \frac{2}{R} . \tag{4.166}$$

Recall that $\mathrm{real}(1/q) = 1/R(z)$ [see Eq. (4.65)], and thus for a cavity mode having a fixed position-dependent radius of curvature $R(z)$ the following must hold

$$R(z_1) = -R_1 , \tag{4.167}$$

$$R(z_2) = R_2 , \tag{4.168}$$

where [see Eq. (4.78)]

$$R(z) = z\left(1 + \frac{|q_0|^2}{z^2}\right) , \tag{4.169}$$

and

$$q_0 = q(z = 0) = \frac{n_0 k_0 w_0^2}{2i} , \tag{4.170}$$

where w_0 is the spot size at the the location of the waist, i.e. at $z = 0$. Equations (4.167) and (4.168) together with the requirement that $z_2 - z_1 = d$ can be expressed as

$$d - \frac{|q_0|^2 d}{z_0^2 - \frac{d^2}{4}} = 2R_0 , \tag{4.171}$$

$$-z_0 - \frac{|q_0|^2 z_0}{z_0^2 - \frac{d^2}{4}} = R_{\mathrm{m}} , \tag{4.172}$$

where

$$R_0 = \frac{R_1 + R_2}{2} , \tag{4.173}$$

$$R_{\mathrm{m}} = \frac{R_1 - R_2}{2} , \tag{4.174}$$

$$z_0 = \frac{z_1 + z_2}{2} , \tag{4.175}$$

and thus

$$z_0 = \frac{d}{2}\frac{R_{\mathrm{m}}}{R_0 - d} , \tag{4.176}$$

and

$$|q_0|^2 = \left(\frac{d}{2}\right)^2 \left(1 + \frac{R_{\mathrm{m}}}{z_0}\right) \left(1 - \frac{R_{\mathrm{m}}^2}{(R_0 - d)^2}\right) , \tag{4.177}$$

or in terms of R_1 and R_2

$$|q_0|^2 = d\frac{(R_1 - d)(R_2 - d)(R_1 + R_2 - d)}{(R_1 + R_2 - 2d)^2} , \tag{4.178}$$

$$z_1 = -\frac{d(R_2 - d)}{R_1 + R_2 - 2d} , \tag{4.179}$$

$$z_2 = \frac{d(R_1 - d)}{R_1 + R_2 - 2d} , \tag{4.180}$$

and thus [see Eq. (4.170)]

$$w_0 = \left[\frac{4d(R_1 - d)(R_2 - d)(R_1 + R_2 - d)}{n_0^2 k_0^2 (R_1 + R_2 - 2d)^2}\right]^{1/4} . \tag{4.181}$$

9. The complex beam parameter at the output q_2 is related to the input parameter q_1 by [see Eqs. (4.10) and (4.66)]

$$q_2 = \frac{q_1}{-\frac{q_1}{f} + 1} , \tag{4.182}$$

hence [see Eq. (4.65)]

$$\frac{1}{R_2} = -\frac{1}{f} + \frac{1}{R_1} . \tag{4.183}$$

10. The solutions of Eq. (4.120) are given by

$$\frac{1}{q} = \frac{D - A \pm \sqrt{D^2 - 2DA + A^2 + 4CB}}{2B} . \tag{4.184}$$

Note that only the solution with the plus sign yields a real waist width w [see Eq. (4.65)]. With the help of Eq. (4.65) one finds that

$$R = \frac{2B}{D - A} . \tag{4.185}$$

11. 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)]

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix} = \begin{pmatrix} 1 & d_{\mathrm{f}} \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ -\frac{1}{f} & 1 \end{pmatrix} = \begin{pmatrix} 1 - \frac{d_{\mathrm{f}}}{f} & d_{\mathrm{f}} \\ -\frac{1}{f} & 1 \end{pmatrix} . \tag{4.186}$$

The following holds [see Eq. (4.66)]

$$q\left(d_{\mathrm{f}}\right) = \frac{Aq_0 + B}{Cq_0 + D} , \tag{4.187}$$

where [see Eqs. (4.65), (4.78) and (4.79) and recall that $k_0 = 2\pi/\lambda$]

$$\frac{1}{q_0} = \frac{i\lambda}{\pi w_0^2} , \tag{4.188}$$

$$\frac{1}{q\left(d_{\mathrm{f}}\right)} = \frac{i\lambda}{\pi w_{\mathrm{f}}^2} , \tag{4.189}$$

and thus Eq. (4.187) becomes

$$\frac{1}{\frac{2ic}{\omega w_{\mathrm{f}}^2}} = \frac{-\frac{1}{f}\left(1 - \frac{d_{\mathrm{f}}}{f}\right) + d_{\mathrm{f}}\left(\frac{\lambda}{\pi w_0^2}\right)^2 - \frac{i\lambda}{\pi w_0^2}}{\frac{1}{f^2} + \left(\frac{\lambda}{\pi w_0^2}\right)^2} . \tag{4.190}$$

The real part of Eq. (4.190) yields

$$d_{\mathrm{f}} = \frac{f}{1+\eta^2} , \tag{4.191}$$

and the imaginary part yields

$$w_{\mathrm{f}}^2 = \frac{w_0^2}{1+\eta^{-2}} , \tag{4.192}$$

where

$$\eta = \frac{f\lambda}{\pi w_0^2} . \tag{4.193}$$

12. The following holds [see Eqs. (4.78) and (4.79)]

$$R_1 = z_1\left(1 + \frac{\left|q_0\right|^2}{z_1^2}\right) , \tag{4.194}$$

and

$$w_1^2 = w_0^2\left(1 + \frac{z_1^2}{\left|q_0\right|^2}\right) , \tag{4.195}$$

where $R = R_1$ and $w = w_1$ at $z = z_1$, the waste is at $z = 0$, q_0 is given by

$$\frac{1}{q_0} = \frac{2i}{k_0 w_0^2} , \tag{4.196}$$

and $k_0 = \omega/c$. The solution for w_0 is given by

$$w_0 = \frac{w_1}{\sqrt{1 + \left(\frac{k_0 w_1^2}{2R_1}\right)^2}} . \tag{4.197}$$

13. The allowed values of the wave propagation coefficient k are found by solving

$$\det\left(M_{\mathrm{B}} - e^{ikL} \times \mathbf{1}\right) = 0 , \tag{4.198}$$

where the transfer matrix M_{B} is given by Eq. (4.115), L is the length of the ring, and $\mathbf{1}$ is the 2×2 identity matrix. The following hold

$$\det M = \frac{1 - |r_{\mathrm{B}}|^2}{|t_{\mathrm{B}}|^2} = 1 , \tag{4.199}$$

and therefore the eigenvalues $\lambda_\pm$ of M_{B} are given by

$$\lambda_\pm = \tau \pm \sqrt{\tau^2 - 1} , \tag{4.200}$$

where

$$\tau = \frac{\operatorname{Tr} M_{\mathrm{B}}}{2} = \frac{1}{2}\left(\frac{1}{t_{\mathrm{B}}} + \frac{1}{t_{\mathrm{B}}^*}\right) = \operatorname{Re}\frac{1}{t_{\mathrm{B}}} , \tag{4.201}$$

hence $\lambda_+\lambda_- = 1$, and $|\lambda_\pm| = 1$ provided that

$$|\tau| \leq 1 . \tag{4.202}$$

In the region where $|\tau| \leq 1$, the eigenvalues can be expressed as $\lambda_\pm = e^{\pm i\theta}$, where the real angle θ is given by

$$\theta = \cos^{-1}\tau . \tag{4.203}$$

For the case of a fiber Bragg grating (FBG) [see Eqs. (4.116) and (4.201)]

$$\tau = \operatorname{Re}\left(\cosh\sqrt{V^2 - \Delta^2} - \frac{i\Delta \sinh\sqrt{V^2 - \Delta^2}}{\sqrt{V^2 - \Delta^2}}\right) . \tag{4.204}$$

In the region where $V^2 - \Delta^2 \leq 0$ [recall that $\cosh(ix) = \cos x$ and $\sinh(ix) = i\sin x$]

$$\tau = \cos\sqrt{\Delta^2 - V^2} , \tag{4.205}$$

hence [see Eq. (4.203)]

$$\theta = \sqrt{\Delta^2 - V^2} . \tag{4.206}$$

14. The set of two coupled first order differential equations for the amplitudes $A_{\pm}$ (4.104), which is given by (it is assumed that $\mathcal{D}$ is real)

$$\frac{\mathrm{d}}{\mathrm{d}\zeta}\begin{pmatrix} A_+ \\ A_- \end{pmatrix} = \begin{pmatrix} i\delta & i\mathcal{D} \\ -i\mathcal{D} & -i\delta \end{pmatrix}\begin{pmatrix} A_+ \\ A_- \end{pmatrix} , \tag{4.207}$$

can be used for deriving a set of two decoupled second order differential equations. This can be done by substituting the transformation

$$\begin{pmatrix} A_+ \\ A_- \end{pmatrix} = \begin{pmatrix} \frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \end{pmatrix}\begin{pmatrix} B_+ \\ B_- \end{pmatrix} \tag{4.208}$$

into Eq. (4.207), which yields

$$\frac{\mathrm{d}}{\mathrm{d}\zeta}\begin{pmatrix} B_+ \\ B_- \end{pmatrix} = \mathcal{M}'\begin{pmatrix} B_+ \\ B_- \end{pmatrix} , \tag{4.209}$$

where

$$\begin{aligned}\mathcal{M}' &= \begin{pmatrix} \frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \end{pmatrix}^{-1}\begin{pmatrix} i\delta & i\mathcal{D} \\ -i\mathcal{D} & -i\delta \end{pmatrix}\begin{pmatrix} \frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \end{pmatrix} \\ &= \begin{pmatrix} 0 & -i\left(\delta - \mathcal{D}\right) \\ -i\left(\delta + \mathcal{D}\right) & 0 \end{pmatrix} .\end{aligned} \tag{4.210}$$

Alternatively, Eq. (4.209) can be rewritten as [see Eq. (4.95), recall that $\zeta = \pi z/\Lambda$ and note that $k_0 = \omega/c \simeq \omega_{\mathrm{B}}/c$ near the Bragg frequency ω_{B}]

$$\frac{\mathrm{d}}{\mathrm{d}z}\begin{pmatrix} B_+ \\ B_- \end{pmatrix} = -ik_0 n_{\mathrm{eff}}\begin{pmatrix} 0 & \delta - \mathcal{D} \\ \delta + \mathcal{D} & 0 \end{pmatrix}\begin{pmatrix} B_+ \\ B_- \end{pmatrix} . \tag{4.211}$$

By applying the derivative $\mathrm{d}/\mathrm{d}z$ to Eq. (4.211) one obtains

$$\left(\frac{\mathrm{d}^2}{\mathrm{d}z^2} + k_0^2 n_{\mathrm{B}}^2\right) B_{\pm} = 0 , \tag{4.212}$$

where the effective longitudinal refractive index n_{B} is given by [see Eq. (4.97) and compare with Eq. (3.3)]

$$n_{\mathrm{B}}^2 = n_{\mathrm{eff}}^2\left[\left(\frac{\omega - \omega_{\mathrm{B}}}{\omega_{\mathrm{B}}}\right)^2 - \mathcal{D}^2\right] . \tag{4.213}$$

The region where n_{B} becomes imaginary, i.e. when $\left(\left(\omega - \omega_{\mathrm{B}}\right)/\omega_{\mathrm{B}}\right)^2 < \mathcal{D}^2$, is commonly referred to as a photonic band gap.

15. The FBG reflection amplitude r_{B} is given by Eq. (4.117), which can be expressed as

$$r_{\mathrm{B}} = \mathcal{R}\left(V, \sqrt{1-\left(\frac{\Delta}{V}\right)^2}\right) , \tag{4.214}$$

where the function $\mathcal{R}(V,S)$ is given by

$$\mathcal{R}(V,S) = \frac{i}{S\coth(VS) - i\sqrt{1-S^2}} , \tag{4.215}$$

V is the grating coupling strength parameter [see Eq. (4.112)] and Δ is the total detuning factor [see Eq. (4.114)]. The reflectivity of an FBG having a relatively large number of periods can be approximated by taking the limit $V \to \infty$ while assuming that the ratio Δ/V remains finite. For the case where $|S| = \sqrt{1-(\Delta/V)^2} \leq 1$, i.e. within the photonic band gap of the FBG, one has

$$\lim_{V\to\infty} \mathcal{R}(V,S) = f(S) , \tag{4.216}$$

where the function $f(S)$ can be expressed using several different forms

$$\begin{aligned} f(S) &= \frac{1}{-\sqrt{1-S^2} - iS} \\ &= -\sqrt{1-S^2} + iS \\ &= \frac{1 + \left(\frac{\sqrt{1-S^2}+1}{\sqrt{1-S^2}-1}\right)^{1/2}}{1 - \left(\frac{\sqrt{1-S^2}+1}{\sqrt{1-S^2}-1}\right)^{1/2}} \\ &= -\exp\left(i\sin^{-1}(-S)\right) \\ &= -\exp\left(-i\cos^{-1}\left(\sqrt{1-S^2}\right)\right) . \end{aligned} \tag{4.217}$$

In this limit of large number of periods the FBG reflection amplitude r_{B} becomes [see Eqs. (4.112) and (4.114)]

$$r_{\mathrm{B}} = -\exp\left(-i\cos^{-1}\frac{\delta}{\mathcal{D}}\right) , \tag{4.218}$$

where $\delta = -(\omega - \omega_{\mathrm{B}})/\omega_{\mathrm{B}}$ is the normalized FBG frequency detuning [see Eq. (4.97)] and where $\mathcal{D}$ is the FBG dimensionless modulation strength [see Eq. (4.94)]. Note that, alternatively, in terms of the FBG effective impedance Γ_{FBG}, which is given by

$$\varGamma_{\mathrm{FBG}} = -\sqrt{\frac{\delta - \mathcal{D}}{\delta + \mathcal{D}}} \,, \tag{4.219}$$

the FBG reflection amplitude r_{B} within the photonic band gap can be expressed as

$$r_{\mathrm{B}} = \frac{\varGamma_{\mathrm{FBG}} - 1}{\varGamma_{\mathrm{FBG}} + 1} \,. \tag{4.220}$$

16. The fixed points of a general Möbius transformation (4.32), i.e. the solutions of

$$z = \frac{Az + B}{Cz + D} \,, \tag{4.221}$$

which are given by

$$z_{\pm} = \frac{T_{\mathrm{M}} \pm \sqrt{T_{\mathrm{M}}^2 - D_{\mathrm{M}}} - D}{C} \,, \tag{4.222}$$

where

$$T_{\mathrm{M}} = \frac{A + D}{2} \,, \tag{4.223}$$

$$D_{\mathrm{M}} = AD - BC \,, \tag{4.224}$$

can be expressed in terms of the eigenvalues $\lambda_{\pm}$ of the matrix

$$M = \begin{pmatrix} A & B \\ C & D \end{pmatrix} \,, \tag{4.225}$$

which are given by

$$\lambda_{\pm} = T_{\mathrm{M}} \pm \sqrt{T_{\mathrm{M}}^2 - D_{\mathrm{M}}} \,, \tag{4.226}$$

as

$$z_{\pm} = \frac{\lambda_{\pm} - D}{C} \,. \tag{4.227}$$

Moreover, the following holds

$$f_{A,B,C,D}\left(z_{\pm} + z_1\right) = z_{\pm} + \frac{D_{\mathrm{M}} z_1}{\lambda_{\pm}^2} + O\left(z_1^2\right) \,. \tag{4.228}$$

For the case where M is given by Eq. (4.121) $D_{\mathrm{M}} = 1$ and the eigenvalues $\lambda_{\pm}$ are given by

$$\lambda_{\pm} = \varLambda_{\pm}\left(1 + \frac{\alpha\beta}{2}\right) \,, \tag{4.229}$$

where

$$\Lambda_{\pm}(x) = x \pm \sqrt{x^2 - 1} \, . \tag{4.230}$$

Note that the following holds

$$\Lambda_{\pm}\left(1 + \frac{\alpha\beta}{2}\right) = 1 \pm \sqrt{\alpha\beta} + O(\alpha\beta) \ . \tag{4.231}$$

The fixed point $z_{\pm}$ is stable provide that $|\lambda_{\pm}| > 1$ [see Eq. (4.228)].

17. With the help of the identity (5.46) one finds that the Fourier transform $E(\omega)$ of $E(t)$ is given by

$$E(\omega) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \mathrm{d}t \, E(t) \, e^{-i\omega t} = \frac{E_0}{\sqrt{2\gamma}} e^{-\frac{(\omega - \omega_{\mathrm{p}})^2}{4\gamma}} \ . \tag{4.232}$$

a) Consider a Möbius transformation mapping the Gaussian pulse variable γ from an initial value γ_{in} to a final value γ_{out} according to

$$\gamma_{\mathrm{out}}^{-1} = \frac{A\gamma_{\mathrm{in}}^{-1} + B}{C\gamma_{\mathrm{in}}^{-1} + D} \, , \tag{4.233}$$

where the parameters A, B, C and D are all constants. For the frequency-like transformation (for which $\gamma_{\mathrm{out}}^{-1} = \gamma_{\mathrm{in}}^{-1} + \gamma_{\mathrm{F}}^{-1}$) the parameters A, B, C and D are given by [see Eqs. (4.124) and (4.232)]

$$M_{\mathrm{F}} = \begin{pmatrix} A & B \\ C & D \end{pmatrix} = \begin{pmatrix} 1 & \gamma_{\mathrm{F}}^{-1} \\ 0 & 1 \end{pmatrix} \, , \tag{4.234}$$

and for the time-like transformation (for which $\gamma_{\mathrm{out}} = \gamma_{\mathrm{in}} + \gamma_{\mathrm{T}}$) the parameters A, B, C and D are given by [see Eq. (4.125)]

$$M_{\mathrm{T}} = \begin{pmatrix} A & B \\ C & D \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ \gamma_{\mathrm{T}} & 1 \end{pmatrix} \, . \tag{4.235}$$

b) The matrix corresponding to a concatenating of a frequency-like M_{F} (4.124) and a time-like M_{T} (4.125) transformations is given by [see Eq. (4.33)]

$$M_0 = M_{\mathrm{F}} M_{\mathrm{T}} = \begin{pmatrix} 1 + \gamma_{\mathrm{T}}\gamma_{\mathrm{F}}^{-1} & \gamma_{\mathrm{F}}^{-1} \\ \gamma_{\mathrm{T}} & 1 \end{pmatrix} \, . \tag{4.236}$$

Note that $\det M_0 = 1$ since $\det M_{\mathrm{F}} = \det M_{\mathrm{T}} = 1$. The fixed points are given by [see Eq. (4.227)]

$$\gamma_{\pm}^{-1} = \frac{\lambda_{\pm} - 1}{\gamma_{\mathrm{T}}} \, , \tag{4.237}$$

where the eigenvalues $\lambda_{\pm}$ are given by [see Eq. (4.229)]

$$\lambda_\pm = 1 + \frac{\gamma_{\mathrm{T}}\gamma_{\mathrm{F}}^{-1}}{2} \pm \sqrt{\left(1 + \frac{\gamma_{\mathrm{T}}\gamma_{\mathrm{F}}^{-1}}{2}\right)^2 - 1}$$
$$= 1 \pm \sqrt{\gamma_{\mathrm{T}}\gamma_{\mathrm{F}}^{-1}} + O\left(\gamma_{\mathrm{T}}\gamma_{\mathrm{F}}^{-1}\right) \ . \tag{4.238}$$

The fixed point $\gamma_\pm$ is stable provide that $|\lambda_\pm| > 1$ [see Eq. (4.228)].

c) The effect of the oscillating mirror on the pulse is described by the transformation [see Eq. (4.122)]

$$E(t) \rightarrow E'(t) = t_{\mathrm{m}}(t) E(t) \ , \tag{4.239}$$

where the phase factor $t_{\mathrm{m}}(t)$ is given by

$$t_{\mathrm{m}}(t) = \exp\left(-\frac{2i\omega_{\mathrm{p}} l_{\mathrm{m}} \sin\left(\omega_{\mathrm{m}}(t - t_0)\right)}{c}\right)$$
$$= \exp\left(i\vartheta_{\mathrm{Fm}} + i\Omega_{\mathrm{Tm}} t - \gamma_{\mathrm{Tm}} t^2\right) + O\left((\omega_{\mathrm{m}} t)^3\right) \ , \tag{4.240}$$

t_0 is the time at which the peak of the pulse hits the mirror, the phase shift ϑ_{Fm} is given by

$$\vartheta_{\mathrm{Fm}} = \frac{2\omega_{\mathrm{p}} l_{\mathrm{m}} \sin\left(\omega_{\mathrm{m}} t_0\right)}{c} \ , \tag{4.241}$$

the Doppler frequency shift Ω_{Tm} is given by

$$\Omega_{\mathrm{Tm}} = -\frac{2\omega_{\mathrm{m}}\omega_{\mathrm{p}} l_{\mathrm{m}} \cos\left(\omega_{\mathrm{m}} t_0\right)}{c} \ , \tag{4.242}$$

and the term $-\gamma_{\mathrm{Tm}} t^2$ gives rise to a linear frequency chirp to the pulse, where the purely imaginary coefficient γ_{Tm} is given by

$$\gamma_{\mathrm{Tm}} = \frac{i\omega_{\mathrm{m}}^2 \omega_{\mathrm{p}} l_{\mathrm{m}} \sin\left(\omega_{\mathrm{m}} t_0\right)}{c} \ . \tag{4.243}$$

Fixed points occur when Ω_{Tm} vanishes, i.e. $|\sin(\omega_{\mathrm{m}} t_0)| = 1$, and thus to lowest nonvanishing order in ω_{m} the stable fixed value of the pulse parameter γ is given by [see Eq. (4.237)]

$$\gamma = \sqrt{\frac{i\omega_{\mathrm{m}}^2 \gamma_{\mathrm{F}} \omega_{\mathrm{p}} l_{\mathrm{m}}}{c}} \ . \tag{4.244}$$

This value corresponds to pulses hitting the mirror when the mirror velocity in the inwards direction peaks.

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 ω. 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 (2.151), which is given by

$$\left(\boldsymbol{\nabla}^2 + k^2\right) u = 0 , \tag{5.1}$$

where

$$k = \frac{\omega}{c} . \tag{5.2}$$

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

$$u(k_x, k_y, z) = \mathcal{F}(u(x, y, z)) , \tag{5.3}$$

where

$$\mathcal{F}(u(x, y, z)) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \mathrm{d}x \mathrm{d}y \; u(x, y, z) \, e^{-i(k_x x + k_y y)} . \tag{5.4}$$

Claim. The inverse Fourier transform is given by

$$\mathcal{F}^{-1}(u(k_x, k_y, z)) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \mathrm{d}k_x \mathrm{d}k_y \; u(k_x, k_y, z) \, e^{i(k_x x + k_y y)} . \tag{5.5}$$

Proof. With the help of the identity

$$\int_{-\infty}^{\infty} \mathrm{d}\kappa \; e^{i\kappa s} = 2\pi\delta(s) \ , \tag{5.6}$$

one obtains [see Eqs. (5.4) and (5.5)]

$$\begin{aligned}
&\mathcal{F}^{-1}\left(\mathcal{F}\left(u\left(x',y',z\right)\right)\right) \\
&= \frac{1}{4\pi^2}\int_{-\infty}^{\infty}\mathrm{d}x\int_{-\infty}^{\infty}\mathrm{d}y\; u\left(x,y,z\right)\int_{-\infty}^{\infty}\mathrm{d}k_x\; e^{ik_x\left(x'-x\right)}\int_{-\infty}^{\infty}\mathrm{d}k_y\; e^{ik_y\left(y'-y\right)} \\
&= u\left(x',y',z\right) \ .
\end{aligned} \tag{5.7}$$

The two-dimensional Fourier transformed function in the plane $z = 0$ is denoted by

$$U\left(k_x,k_y\right) = u\left(k_x,k_y,z=0\right) \ , \tag{5.8}$$

where [see Eq. (5.4)]

$$U\left(k_x,k_y\right) = \frac{1}{2\pi}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\mathrm{d}x\mathrm{d}y\; u\left(x,y,0\right)e^{-i\left(k_x x+k_y y\right)} \ . \tag{5.9}$$

Claim. The following holds

$$u\left(k_x,k_y,z\right) = U\left(k_x,k_y\right)e^{ik_z z} \ , \tag{5.10}$$

where

$$k_z = \sqrt{k^2 - k_x^2 - k_y^2} \ . \tag{5.11}$$

Proof. By substituting $u\left(x,y,z\right) = \mathcal{F}^{-1}\left(u\left(k_x,k_y,z\right)\right)$ [see Eq. (5.5)] into the Helmholtz equation (5.1) one obtains

$$\frac{1}{2\pi}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\mathrm{d}k_x\mathrm{d}k_y\; e^{i\left(k_x x+k_y y\right)}\left(k_z^2+\frac{\partial^2}{\partial z^2}\right)u\left(k_x,k_y,z\right) = 0 \ , \tag{5.12}$$

and thus

$$\left(k_z^2+\frac{\partial^2}{\partial z^2}\right)u\left(k_x,k_y,z\right) = 0 \ . \tag{5.13}$$

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 spacial frequency, for which $k_x^2 + k_y^2 > k^2$, do no 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

$$U(k_x, k_y) = \frac{1}{2\pi i k_z} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \mathrm{d}x \mathrm{d}y \left. \frac{\partial u}{\partial z} \right|_{z=0} e^{-i(k_x x + k_y y)} . \tag{5.14}$$

Proof. The following holds [see Eqs. (5.5) and (5.10)]

$$u(x, y, z) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \mathrm{d}k_x \mathrm{d}k_y \, U(k_x, k_y) \, e^{i\mathbf{k}\cdot\mathbf{r}} , \tag{5.15}$$

where $\mathbf{r} = (x, y, z)$, $\mathbf{k} = (k_x, k_y, k_z)$, and k_z is given by Eq. (5.11). Taking the derivative with respect to z and setting $z = 0$ lead to

$$\left. \frac{\partial u}{\partial z} \right|_{z=0} = \frac{1}{2\pi} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \mathrm{d}k_x \mathrm{d}k_y \, i k_z U(k_x, k_y) \, e^{i(k_x x + k_y y)} . \tag{5.16}$$

The expression given by Eq. (5.14) is obtained by applying the two-dimensional Fourier transform (5.4), i.e. by multiplying by $e^{-i\left(k_x' x + k_y' y\right)}$, 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

$$u(\mathbf{r}') = u_1(\mathbf{r}') = u_2(\mathbf{r}') , \tag{5.17}$$

where

$$u_1(\mathbf{r}') = 2 \int_{-\infty}^{\infty} \mathrm{d}x'' \int_{-\infty}^{\infty} \mathrm{d}y'' \, u(\mathbf{r}'') \frac{\partial g(\mathbf{r}' - \mathbf{r}'')}{\partial z''} , \tag{5.18}$$

$$u_2(\mathbf{r}') = -2 \int_{-\infty}^{\infty} \mathrm{d}x'' \int_{-\infty}^{\infty} \mathrm{d}y'' \, \frac{\partial u(\mathbf{r}'')}{\partial z''} g(\mathbf{r}' - \mathbf{r}'') \ , \tag{5.19}$$

where $\mathbf{r}' = (x', y', z')$, $\mathbf{r}'' = (x'', y'', z'')$, in both Eqs. (5.18) and (5.19) the integrals are evaluated in the plane $z'' = 0$, the function g is given by

$$g(\mathbf{r}) = \frac{i}{8\pi^2} \int_{-\infty}^{\infty} \mathrm{d}k_x \int_{-\infty}^{\infty} \mathrm{d}k_y \, \frac{e^{i\mathbf{k}\cdot\mathbf{r}}}{k_z} \ , \tag{5.20}$$

and k_z is given by Eq. (5.11).

Proof. With the help of Eqs. (5.9) and (5.15) one finds that

$$\begin{aligned} u(\mathbf{r}') &= \frac{1}{4\pi^2} \int_{-\infty}^{\infty} \mathrm{d}x'' \int_{-\infty}^{\infty} \mathrm{d}y'' \int_{-\infty}^{\infty} \mathrm{d}k_x \int_{-\infty}^{\infty} \mathrm{d}k_y \; u(x'', y'', 0) \, e^{i\mathbf{k}\cdot\mathbf{r}' - i\left(k_x x'' + k_y y''\right)} \\ &= \frac{1}{4\pi^2} \int_{-\infty}^{\infty} \mathrm{d}x'' \int_{-\infty}^{\infty} \mathrm{d}y'' \; u(\mathbf{r}'') \int_{-\infty}^{\infty} \mathrm{d}k_x \int_{-\infty}^{\infty} \mathrm{d}k_y e^{i\mathbf{k}\cdot\left(\mathbf{r}' - \mathbf{r}''\right)} \\ &= 2 \int_{-\infty}^{\infty} \mathrm{d}x'' \int_{-\infty}^{\infty} \mathrm{d}y'' \; u(\mathbf{r}'') \times \frac{\partial g(\mathbf{r}' - \mathbf{r}'')}{\partial z''} \ . \end{aligned} \tag{5.21}$$

Similarly, Eqs. (5.14) and (5.15) yield

$$\begin{aligned} u(\mathbf{r}') &= \frac{1}{4\pi^2} \int_{-\infty}^{\infty} \mathrm{d}x'' \int_{-\infty}^{\infty} \mathrm{d}y'' \; \left.\frac{\partial u}{\partial z''}\right|_{z''=0} \int_{-\infty}^{\infty} \mathrm{d}k_x \int_{-\infty}^{\infty} \mathrm{d}k_y \; \frac{e^{i\mathbf{k}\cdot\mathbf{r}' - i\left(k_x x'' + k_y y''\right)}}{i k_z} \\ &= -2 \int_{-\infty}^{\infty} \mathrm{d}x'' \int_{-\infty}^{\infty} \mathrm{d}y'' \; \left.\frac{\partial u}{\partial z''}\right|_{z''=0} g(\mathbf{r}' - \mathbf{r}'') \ . \end{aligned} \tag{5.22}$$

The expression for $u_1(\mathbf{r}')$ (5.18) and $u_2(\mathbf{r}')$ (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 $g(\mathbf{r})$ [see Eq. (5.20)] can be expressed as

$$g(\mathbf{r}) = u_{\mathrm{S}}(\mathbf{r}) \ , \tag{5.23}$$

where the scalar spherical wave $u_{\mathrm{S}}(\mathbf{r})$ is defined by

$$u_{\mathrm{S}}(\mathbf{r}) = -\frac{e^{ikr}}{4\pi r} \ , \tag{5.24}$$

k is given by Eq. (5.2) and $r = \sqrt{x^2 + y^2 + z^2}$.

Proof. The two-dimensional Fourier transform of $u_{\mathrm{S}}(\mathbf{r})$ in the plane $z=0$, which is denoted by $U_{\mathrm{S}}(k_x,k_y)$, is given by [see Eq. (5.9)]

$$\begin{aligned} U_{\mathrm{S}}(k_x,k_y) &= \frac{1}{2\pi}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \mathrm{d}x\mathrm{d}y\; u_{\mathrm{S}}(x,y,0)\, e^{-i(k_x x+k_y y)} \\ &= -\frac{1}{8\pi^2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \mathrm{d}x\mathrm{d}y\; \frac{e^{ik\sqrt{x^2+y^2}}}{\sqrt{x^2+y^2}} e^{-i(k_x x+k_y y)}\ . \end{aligned} \tag{5.25}$$

The coordinate transformation

$$x = \rho\cos\theta\ , \tag{5.26}$$

$$y = \rho\sin\theta\ , \tag{5.27}$$

$$k_x = \kappa\cos\varphi\ , \tag{5.28}$$

$$k_y = \kappa\sin\varphi\ , \tag{5.29}$$

together with the identity

$$\frac{1}{2\pi}\int_{-\pi}^{\pi}\mathrm{d}\theta e^{-iA\cos\theta} = J_0(A)\ , \tag{5.30}$$

where J_0 is the Bessel function of the first kind, yield

$$\begin{aligned} U_{\mathrm{S}}(k_x,k_y) &= -\frac{1}{8\pi^2}\int_{-\pi}^{\pi}\mathrm{d}\theta\int_{0}^{\infty}\mathrm{d}\rho\;\rho\frac{e^{ik\rho}}{\rho}e^{-i\kappa\rho\cos(\varphi-\theta)} \\ &= -\frac{1}{4\pi}\int_{0}^{\infty}\mathrm{d}\rho\; e^{ik\rho}J_0(\kappa\rho)\ . \end{aligned} \tag{5.31}$$

The identity

$$\int_{0}^{\infty}\mathrm{d}\rho\; e^{ik\rho}J_0(\kappa\rho) = \frac{1}{i\sqrt{k^2-\kappa^2}}\ , \tag{5.32}$$

leads to

$$U_{\mathrm{S}}(k_x,k_y) = -\frac{1}{4\pi i k_z}\ , \tag{5.33}$$

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

$$\frac{e^{ikr}}{r} = \frac{1}{2\pi i} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \mathrm{d}k_x \mathrm{d}k_y \, \frac{e^{i\mathbf{k}\cdot\mathbf{r}}}{k_z} \; , \tag{5.34}$$

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

$$u_1(\mathbf{r}') = -\frac{1}{2\pi} \int_{-\infty}^{\infty} \mathrm{d}x'' \int_{-\infty}^{\infty} \mathrm{d}y'' \, u(\mathbf{r}'') \frac{\partial}{\partial z''} \left(\frac{e^{ikr}}{r} \right) \; , \tag{5.35}$$

whereas the so-called second Rayleigh-Sommerfeld diffraction integral is obtained by combining Eqs. (5.19) and (5.23)

$$u_2(\mathbf{r}') = \frac{1}{2\pi} \int_{-\infty}^{\infty} \mathrm{d}x'' \int_{-\infty}^{\infty} \mathrm{d}y'' \, \frac{\partial u(\mathbf{r}'')}{\partial z''} \frac{e^{ikr}}{r} \; , \tag{5.36}$$

where

$$r = |\mathbf{r}' - \mathbf{r}''| \; . \tag{5.37}$$

The Kirchhoff diffraction integral $u_{\mathrm{K}}(\mathbf{r}')$ is defined by [see Eqs. (5.18) and (5.19) and compare with Eq. (5.85) below]

$$\begin{aligned} u_{\mathrm{K}}(\mathbf{r}') &= \frac{u_1(\mathbf{r}') + u_2(\mathbf{r}')}{2} \\ &= \int_{-\infty}^{\infty} \mathrm{d}x'' \int_{-\infty}^{\infty} \mathrm{d}y'' \left(u(\mathbf{r}'') \frac{\partial g(\mathbf{r}' - \mathbf{r}'')}{\partial z''} - \frac{\partial u(\mathbf{r}'')}{\partial z''} g(\mathbf{r}' - \mathbf{r}'') \right) \; . \end{aligned} \tag{5.38}$$

5.2.1 The Limit of Geometrical Optics

Consider in general an integral I, which is given by

$$I = \int_{-\infty}^{\infty} \mathrm{d}x \; h(x) \, e^{ikf(x)} \; , \tag{5.39}$$

where the functions $h(x)$ and $f(x)$ and the coefficient k are all real. Let $I_{\mathrm{d}}(x_0, \delta)$ be the contribution to I from the interval $[x_0 - \delta, x_0 + \delta]$, i.e.

$$I_{\mathrm{d}}(x_0,\delta)=\int_{x_0-\delta}^{x_0+\delta}\mathrm{d}x\;h(x)\,e^{ikf(x)}\;. \tag{5.40}$$

For sufficiently small δ the following approximations are expected to hold inside the interval $[x_0-\delta,x_0+\delta]$

$$h(x)\simeq h(x_0)\;, \tag{5.41}$$

$$f(x)\simeq f(x_0)+(x-x_0)\,f'(x_0)\;, \tag{5.42}$$

and thus

$$I_{\mathrm{d}}(x_0,\delta)\simeq 2h(x_0)\,e^{ikf(x_0)}\frac{\sin(\delta kf'(x_0))}{kf'(x_0)}\;. \tag{5.43}$$

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 $kf(x)$ in Eq. (5.39) is locally stationary, i.e. points x_{s} such that $f'(x_{\mathrm{s}})=0$.

Exercise 5.2.1. Calculate I in the limit $k\rightarrow\infty$ 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}

$$f(x)=f(x_{\mathrm{s}})+\frac{(x-x_{\mathrm{s}})^2}{2}f''(x_{\mathrm{s}})+\cdots\;, \tag{5.44}$$

where prime denotes a derivative with respect to x, and the variable transformation $s=x-x_{\mathrm{s}}$, one obtains in the limit $k\rightarrow\infty$

$$I\simeq e^{ikf(x_{\mathrm{s}})}h(x_{\mathrm{s}})\int_{-\infty}^{\infty}\mathrm{d}s\;e^{ik\frac{s^2}{2}f''(x_{\mathrm{s}})}\;. \tag{5.45}$$

With the help of the identity

$$\int_{-\infty}^{\infty}\mathrm{d}x\;\exp\left(-ax^2+bx+c\right)=\sqrt{\frac{\pi}{a}}e^{\frac{1}{4}\frac{4ca+b^2}{a}}\;, \tag{5.46}$$

this becomes

$$I\simeq h(x_{\mathrm{s}})\sqrt{\frac{2\pi i}{kf''(x_{\mathrm{s}})}}e^{ikf(x_{\mathrm{s}})}\;. \tag{5.47}$$

The above result (5.47) is know as the stationary phase approximation. The characteristic width Δx 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

$$\Delta x=\frac{1}{\sqrt{kf''(x_{\mathrm{s}})}}\;. \tag{5.48}$$

Exercise 5.2.2. Calculate the Rayleigh-Sommerfeld first diffraction integral (5.35) in the limit $k \rightarrow \infty$.

Solution 5.2.2. With the help of Eqs. (5.35) and (5.53) one obtains

$$u_1(\mathbf{r}') = \frac{1}{2\pi} \int_{-\infty}^{\infty} \mathrm{d}x'' \int_{-\infty}^{\infty} \mathrm{d}y''\, u(\mathbf{r}'') \left(\frac{e^{ikr}(ikr-1)\, z'}{r^3} \right) , \tag{5.49}$$

where

$$r = z' \sqrt{1 + \frac{(x'-x'')^2 + (y'-y'')^2}{z'^2}} , \tag{5.50}$$

thus in the limit $k \rightarrow \infty$ [see Eq. (5.47)]

$$u_1(\mathbf{r}') = u(x', y', 0)\, e^{ikz'} . \tag{5.51}$$

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 $u(\mathbf{r}'')$ 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 $r \simeq z'$, 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

$$u_1(\mathbf{r}') = \frac{ike^{ikz'}}{2\pi z'} \int_{-\infty}^{\infty} \mathrm{d}x'' \int_{-\infty}^{\infty} \mathrm{d}y''\ u(\mathbf{r}'')\, e^{ik\frac{(x'-x'')^2+(y'-y'')^2}{2z'}} \ . \tag{5.56}$$

Exercise 5.3.1. Consider the case where the scalar $u(\mathbf{r}'')$ in the plane $z'' = 0^+$ is taken to be given by

$$u\left(x'', y'', z'' = 0^+\right) = u_{\mathrm{inc}}(x'', y'')\, t(x'', y'') \ , \tag{5.57}$$

where for normal incident $u_{\mathrm{inc}}(x'', y'')$ is assumed to be a positive constant denoted by u_0 and the aperture transmission $t(x'', y'')$ is given by

$$t(x'', y'') = \begin{cases} 1 & x'' < 0 \\ 0 & x'' \geq 0 \end{cases} \ . \tag{5.58}$$

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 $|u_1(\mathbf{r}')| \ll u_0$).

Solution 5.3.1. In general, the Fresnel diffraction integral (5.56) together with Eq. (5.48) imply that the value of $u_1(\mathbf{r}')$ 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 $\Delta x'' = \Delta y'' = \sqrt{z'/k}$. For the current case under consideration $u(x'', y'', z'' = 0^+)$ vanishes when $x'' \geq 0$, and consequently it is expected that $|u_1(\mathbf{r}')| \ll u_0$ provided that $x'' \gtrsim \Delta x'' = \sqrt{z'/k}$.

Fraunhofer diffraction integral [see Eq. (5.60) below] is derived from the additional assumption that

$$\frac{k\left(x''^2 + y''^2\right)}{2z'} \ll 1 \ . \tag{5.59}$$

In this limit Eq. (5.56) leads to the Fraunhofer diffraction integral

$$u_1(\mathbf{r}') = \frac{ik\Phi}{2\pi z'} \int_{-\infty}^{\infty} \mathrm{d}x'' \int_{-\infty}^{\infty} \mathrm{d}y''\ u(\mathbf{r}'')\, e^{-i\left(\kappa_x x'' + \kappa_y y''\right)} \ , \tag{5.60}$$

where

$$\Phi = e^{ik\left(z' + \frac{x'^2 + y'^2}{2z'}\right)} \ , \tag{5.61}$$

and where

$$\kappa_x = \frac{kx'}{z'} \ , \kappa_y = \frac{ky'}{z'} \ . \tag{5.62}$$

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 $u(\mathbf{r}'')$.

5.4 Imaging

Consider a lens having transmission $t_{\mathrm{L}}(x,y)$, which is taken to be given by

$$t_{\mathrm{L}}(x,y)=e^{-\frac{ik\left(x^2+y^2\right)}{2f}}P_{\mathrm{L}}(x,y)\ , \tag{5.63}$$

where f is the focal length of the lens. The so-called pupil function $P_{\mathrm{L}}(x,y)$, which for the present case is taken to be given by

$$P_{\mathrm{L}}(x,y)=\begin{cases}1 & |x|\leq A_{\mathrm{L}} \text{ and } |y|\leq A_{\mathrm{L}} \\ 0 & \text{otherwise}\end{cases}\ , \tag{5.64}$$

accounts for the finite size of the lens. The lens is placed in the plane $z=0$. An aperture having transmission $t_{\mathrm{A}}(x,y)$, which is taken to be given by

$$t_{\mathrm{A}}(x'',y'')=\delta(x''-x_0)\,\delta(y''-y_0)\ , \tag{5.65}$$

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 $P_{\mathrm{L}}(x,y)$ and a general aperture transmission $t_{\mathrm{A}}(x'',y'')$

$$\begin{aligned}u_1(x''',y''',s_2)&=-\frac{k^2e^{ik(s_1+s_2)}u_0}{4\pi^2s_1s_2}\\&\quad\times\int_{-\infty}^{\infty}\mathrm{d}x'\int_{-\infty}^{\infty}\mathrm{d}y'\int_{-\infty}^{\infty}\mathrm{d}x''\int_{-\infty}^{\infty}\mathrm{d}y''\ P_{\mathrm{L}}(x',y')\,t_{\mathrm{A}}(x'',y'')\,e^{\frac{ik\Phi}{2}}\ ,\end{aligned} \tag{5.66}$$

where

$$\begin{aligned}\Phi&=\frac{(x'-x'')^2+(y'-y'')^2}{s_1}+\frac{(x'''-x')^2+(y'''-y')^2}{s_2}-\frac{x'^2+y'^2}{f}\\&=\left(x'^2+y'^2\right)\left(\frac{1}{s_1}+\frac{1}{s_2}-\frac{1}{f}\right)\\&\quad+\frac{x''^2+y''^2}{s_1}+\frac{x'''^2+y'''^2}{s_2}-2x'\left(\frac{x''}{s_1}+\frac{x'''}{s_2}\right)-2y'\left(\frac{y''}{s_1}+\frac{y'''}{s_2}\right)\ .\end{aligned} \tag{5.67}$$

When $P_{\mathrm{L}}(x,y)$ is given by Eq. (5.64) and $t_{\mathrm{A}}(x'',y'')$ by Eq. (5.65) this becomes

$$u_1(x''', y''', s_2) = -\frac{k^2 e^{ik(s_1+s_2)} u_0}{4\pi^2 s_1 s_2} \int_{-A_L}^{A_L} dx' \int_{-A_L}^{A_L} dy' \, e^{\frac{ik\Phi}{2}} , \tag{5.68}$$

where

$$\begin{aligned} \Phi &= \left(x'^2 + y'^2\right)\left(\frac{1}{s_1} + \frac{1}{s_2} - \frac{1}{f}\right) \\ &\quad + \frac{x_0^2 + y_0^2}{s_1} + \frac{x'''^2 + y'''^2}{s_2} - \frac{2x'\left(x''' - Mx_0\right)}{s_2} - \frac{2y'\left(y''' - My_0\right)}{s_2} , \end{aligned} \tag{5.69}$$

and where M, which is given by

$$M = -\frac{s_2}{s_1} , \tag{5.70}$$

is the magnification [compare with Eq. (4.129)].

When the imaging condition, which is given by [compare with Eq. (4.127)]

$$\frac{1}{s_1} + \frac{1}{s_2} = \frac{1}{f} , \tag{5.71}$$

is satisfied the scalar $u_1(x''', y''', s_2)$ in the object plane is found to be given by

$$\begin{aligned} & u_1(x''', y''', s_2) \\ &= -\frac{k^2 e^{ik(s_1+s_2)} e^{ik\left(\frac{x_0^2+y_0^2}{2s_1} + \frac{x'''^2+y'''^2}{2s_2}\right)} u_0}{4\pi^2 s_1 s_2} \\ &\quad \times \int_{-A_L}^{A_L} dx' \, e^{-i\frac{k(x'''-Mx_0)x'}{s_2}} \int_{-A_L}^{A_L} dy' \, e^{-i\frac{k(y'''-My_0)y'}{s_2}} \\ &= -\frac{k^2 A_L^2 e^{ik(s_1+s_2)} e^{ik\left(\frac{x_0^2+y_0^2}{2s_1} + \frac{x'''^2+y'''^2}{2s_2}\right)} u_0}{\pi^2 s_1 s_2} \\ &\quad \times \operatorname{sinc} \frac{k\left(x''' - Mx_0\right) A_L}{s_2} \operatorname{sinc} \frac{k\left(y''' - My_0\right) A_L}{s_2} , \end{aligned} \tag{5.72}$$

where

$$\operatorname{sinc} q = \frac{\sin q}{q} . \tag{5.73}$$

As can be seen from Eq. (5.6), the following holds

$$2\pi\delta(K) = \lim_{A\to\infty}\int_{-A}^{A} \mathrm{d}X'\, e^{iKX'} = 2\lim_{A\to\infty} A\,\mathrm{sinc}(KA)\ , \tag{5.74}$$

thus in the limit where

$$\frac{kx''' A_{\mathrm{L}}}{s_2} \gg 1\ , \tag{5.75}$$

$$\frac{ky''' A_{\mathrm{L}}}{s_2} \gg 1\ , \tag{5.76}$$

$u_1(x''', y''', s_2)$ becomes

$$\begin{aligned} &u_1(x''', y''', s_2)\\ &= -\frac{k^2 e^{ik(s_1+s_2)} e^{ik\left(\frac{x_0^2+y_0^2}{2s_1}+\frac{x'''^2+y'''^2}{2s_2}\right)} u_0}{s_1 s_2}\\ &\quad\times\delta\left(\frac{k(x'''-Mx_0)A_{\mathrm{L}}}{s_2}\right)\delta\left(\frac{k(y'''-My_0)A_{\mathrm{L}}}{s_2}\right)\ , \end{aligned} \tag{5.77}$$

or

$$\begin{aligned} u_1(x''', y''', s_2) = {} & \frac{M e^{ik(s_1+s_2)} e^{ik\left(\frac{x_0^2+y_0^2}{2s_1}+\frac{x'''^2+y'''^2}{2s_2}\right)} u_0}{A_{\mathrm{L}}^2}\\ & \times\delta(x'''-Mx_0)\,\delta(y'''-My_0)\ . \end{aligned} \tag{5.78}$$

Exercise 5.4.2. Calculate $u_1(x''', y''', s_2)$ for the case of a general aperture transmission $t_{\mathrm{A}}(x'', y'')$ 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

$$\begin{aligned} u_1(x''', y''', s_2) = {} & -\frac{k^2 e^{ik(s_1+s_2)} e^{ik\frac{x'''^2+y'''^2}{2f}} u_0}{4\pi^2 s_2}\int_{-A_{\mathrm{L}}}^{A_{\mathrm{L}}} \mathrm{d}x' \int_{-A_{\mathrm{L}}}^{A_{\mathrm{L}}} \mathrm{d}y' e^{-ik\left(\frac{x'''x'+y'''y'}{f}\right)}\\ & \times\lim_{s_1\to 0}\frac{1}{s_1}\int_{-\infty}^{\infty}\mathrm{d}x'' \int_{-\infty}^{\infty}\mathrm{d}y''\ t_{\mathrm{A}}(x'', y'')\, e^{ik\frac{(x'-x'')^2+(y'-y'')^2}{2s_1}}\ , \end{aligned} \tag{5.79}$$

thus [see Eq. (5.47)]

$$u_1(x''', y''', s_2) = -\frac{ike^{ik(s_1+s_2)} e^{ik\frac{x'''^2+y'''^2}{2f}} u_0}{s_2} t_{\mathrm{A}}\left(\frac{kx'''}{f}, \frac{ky'''}{f}; A_{\mathrm{L}}\right)\ . \tag{5.80}$$

where

$$t_{\rm A}\left(k_x,k_y;A_{\rm L}\right)=\frac{1}{2\pi}\int_{-A_{\rm L}}^{A_{\rm L}}{\rm d}x'\int_{-A_{\rm L}}^{A_{\rm L}}{\rm d}y'\;t_{\rm A}\left(x',y'\right)e^{-i\left(k_xx'+k_yy'\right)}\;. \tag{5.81}$$

Note that in the limit where $A_{\rm L}\rightarrow\infty$ the term $t_{\rm A}\left(k_x,k_y;A_{\rm L}\right)$ (5.81) becomes the two-dimensional Fourier transform of the aperture transmission $t_{\rm A}\left(x',y'\right)$ [see Eq. (5.4)].

5.5 Problems

1. **Green's theorem** - Show that

$$\int_V\left(u\boldsymbol{\nabla}^2g-g\boldsymbol{\nabla}^2u\right){\rm d}v=\int_S\left(u\frac{\partial g}{\partial n}-g\frac{\partial u}{\partial n}\right){\rm d}s\;, \tag{5.82}$$

where S is the boundary of the volume V, both $u\left(\mathbf{r}\right)$ and $g\left(\mathbf{r}\right)$ are smooth functions from $\mathcal{R}^3$ to $\mathcal{C}$ and $\partial/\partial n$ denotes a partial derivative in the outward normal direction on the boundary S, i.e.

$$\frac{\partial\psi}{\partial n}=\hat{\mathbf{n}}\cdot\boldsymbol{\nabla}u\;, \tag{5.83}$$

where $\hat{\mathbf{n}}$ is a unit vector normal to the boundary S.

2. Show that

$$\left(\boldsymbol{\nabla}^2+k^2\right)g\left(\mathbf{r}\right)=\delta\left(\mathbf{r}\right)\;, \tag{5.84}$$

where $g\left(\mathbf{r}\right)$ is given by Eq. (5.20).

3. **Kirchhoff diffraction integral** - Let $u\left(\mathbf{r}\right)$ be a solution of the Helmholtz equation (5.1) in a volume V, which is bounded by the surface S. Show that

$$u\left(\mathbf{r}\right)=\frac{1}{4\pi}\int_S\left(\frac{\partial u}{\partial n}\frac{e^{ikr}}{r}-u\frac{\partial}{\partial n}\left(\frac{e^{ikr}}{r}\right)\right){\rm d}s\;, \tag{5.85}$$

where $\partial/\partial n$ denotes a partial derivative in the outward normal direction, $r=\left|\mathbf{r}-\mathbf{r}'\right|$ and $\mathbf{r}'$ 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.

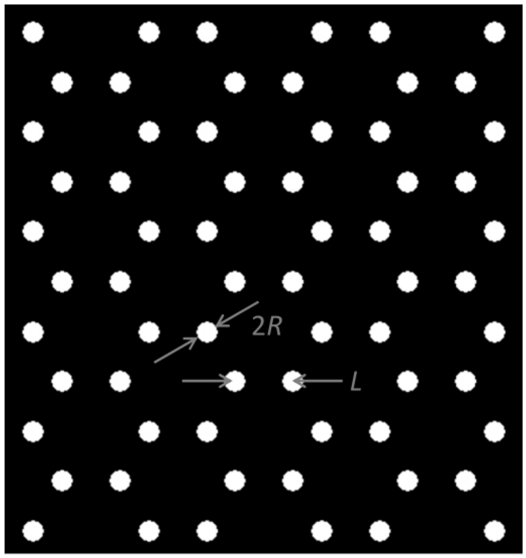

Fig. 5.1. Honeycomb array of holes.

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

$$t(x'', y'') = \frac{1 + \eta \cos \frac{2\pi x''}{a}}{2} , \tag{5.86}$$

where both η 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 λ 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

$$t(x'', y'') = T_{\mathrm{F}}\left(\sqrt{x''^2 + y''^2}\right) , \tag{5.87}$$

where the function T_{F} is given by

$$T_{\mathrm{F}}(r) = \frac{1}{2}\left[1 + \operatorname{sgn}\left(\cos\left(\gamma r^2\right)\right)\right] \;, \tag{5.88}$$

the function sgn is the sign function, i.e.

$$\operatorname{sgn}(\theta) = \begin{cases} 1 & \theta \geq 0 \\ -1 & \theta < 0 \end{cases} \;, \tag{5.89}$$

and γ 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 λ. 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 is 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 $I_m = P_m/P_{\mathrm{in}}$ that is delivered to the m'th focal point.

9. Consider an isosceles triangle aperture normally illuminated by an incident monochromatic plane wave. Calculate the Fraunhofer diffraction integral (5.60). Assume that in the aperture plane the vertices of the triangle are located at the points $\bar{a} = L\left(1/2, \sqrt{3}/2\right)$, $\bar{b} = L\left(-1/2, \sqrt{3}/2\right)$ and $(0, 0)$.
10. Calculate the diffraction efficiency into the first diffraction order for a grating having transmission $t(x'', y'')$

    a) given by

    $$t(x'', y'') = \left|\cos\frac{\pi x''}{L}\right| \;, \tag{5.90}$$

    where L is a constant.

    b) given by

    $$t(x'', y'') = e^{i\phi(x'')} \;, \tag{5.91}$$

    where $\phi(x'')$ is periodic $\phi(x'' + L) = \phi(x'')$ and $\phi(x'') = \phi_0 x''/L$ for $L/2 \leq x'' < L/2$, where both ϕ_0 and L are constants.
11. Calculate the Fraunhofer diffraction integral (5.60) for the case of normal incident and the rectangular aperture seen in Fig. 5.2, having inner a_1 and outer a_2 sides.

5.6 Solutions

1. The following holds [see Eq. (2.149)]

$$\boldsymbol{\nabla}\cdot(g\boldsymbol{\nabla}u) = g\boldsymbol{\nabla}^2 u + \boldsymbol{\nabla}u\cdot\boldsymbol{\nabla}g \;, \tag{5.92}$$

$$\boldsymbol{\nabla}\cdot(u\boldsymbol{\nabla}g) = u\boldsymbol{\nabla}^2 g + \boldsymbol{\nabla}u\cdot\boldsymbol{\nabla}g \;, \tag{5.93}$$

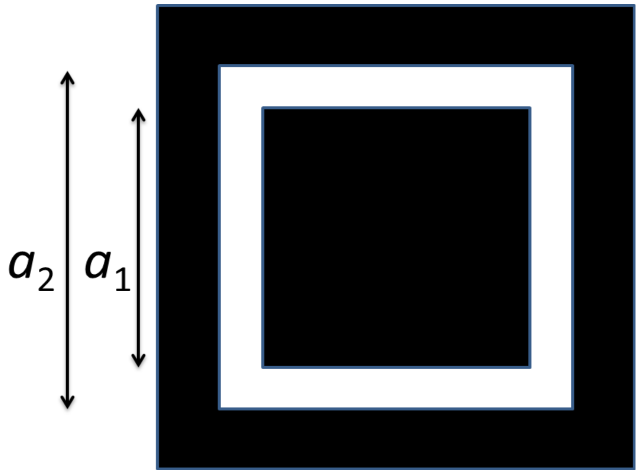

Fig. 5.2. Rectangular frame aperture.

hence

$$u\boldsymbol{\nabla}^2 g - g\boldsymbol{\nabla}^2 u = \boldsymbol{\nabla}\cdot\left(u\boldsymbol{\nabla}g - g\boldsymbol{\nabla}u\right) . \tag{5.94}$$

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

$$g(\mathbf{r}) = -\frac{e^{ikr}}{4\pi r} , \tag{5.95}$$

where $r = |\mathbf{r}|$. The following holds

$$\boldsymbol{\nabla}^2 g = \frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial g}{\partial r}\right) = \frac{k^2 e^{ikr}}{4\pi r} , \tag{5.96}$$

and thus for $r \neq 0$

$$\left(\boldsymbol{\nabla}^2 + k^2\right) g(\mathbf{r}) = 0 . \tag{5.97}$$

Consider a sphere of radius r_0 centered at the origin. The integral of $\boldsymbol{\nabla}^2 g$ 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

$$\int_V \boldsymbol{\nabla}^2 g \,\mathrm{d}v = \int_S \boldsymbol{\nabla}g\cdot \mathrm{d}\mathbf{s} = \int_S e^{ikr}\frac{1-ikr}{4\pi r^2}\hat{\mathbf{r}}\cdot \mathrm{d}\mathbf{s} . \tag{5.98}$$

The volume integral over $k^2 g$ is given by

$$\int_V k^2 g \, \mathrm{d}v = -k^2 \int_0^{r_0} \mathrm{d}r \, r^2 \frac{e^{ikr}}{r} \; . \tag{5.99}$$

Thus, in the limit $r_0 \to 0$

$$\lim_{r_0 \to 0} \int_V \left(\boldsymbol{\nabla}^2 + k^2 \right) g \, \mathrm{d}v = 1 \; , \tag{5.100}$$

and hence Eq. (5.84) holds.

3. By employing the Green's theorem (5.82) with the functions $u(\mathbf{r})$ and $g(\mathbf{r} - \mathbf{r}')$, which is given by Eq. (5.20), and by making use of Eqs. (5.1) and (5.84) one obtains

$$\begin{aligned} \int_S \left(u \frac{\partial g}{\partial n} - g \frac{\partial u}{\partial n} \right) \mathrm{d}s &= \int_V \left(u \boldsymbol{\nabla}^2 g - g \boldsymbol{\nabla}^2 u \right) \mathrm{d}v \\ &= \int_V \left(u \left(\boldsymbol{\nabla}^2 + k^2 \right) g - g \left(\boldsymbol{\nabla}^2 + k^2 \right) u \right) \mathrm{d}v \\ &= u(\mathbf{r}) \; . \end{aligned} \tag{5.101}$$

The above result and the relation (5.23) lead to Eq. (5.85).

4. The scalar $u(\mathbf{r}'')$ in the plane $z'' = 0^+$ is taken to be given by

$$u\left(x'', y'', z'' = 0^+\right) = u_{\mathrm{inc}}(x'', y'') \, t(x'', y'') \; , \tag{5.102}$$

where for normal incident $u_{\mathrm{inc}}(x'', y'')$ is a constant denoted by u_0 and the aperture transmission is given by

$$t(x'', y'') = \begin{cases} 1 & \rho'' \leq a \\ 0 & \rho'' > a \end{cases} \; , \tag{5.103}$$

and

$$\rho'' = \sqrt{x''^2 + y''^2} \; . \tag{5.104}$$

In cylindrical coordinates

$$x'' = \rho'' \cos\theta'' \; , \tag{5.105}$$
$$y'' = \rho'' \sin\theta'' \; , \tag{5.106}$$
$$x' = \rho' \cos\theta' \; , \tag{5.107}$$
$$y' = \rho' \sin\theta' \; , \tag{5.108}$$

the Fresnel diffraction integral (5.56) becomes [see Eq. (5.30)]

$$\begin{aligned}
u_1(\mathbf{r}') &= \frac{ike^{ikz'}u_0}{2\pi z'}\int\limits_0^a \mathrm{d}\rho''\ \rho''\int\limits_{-\pi}^{\pi}\mathrm{d}\theta''\ e^{ik\frac{(\rho'\cos\theta'-\rho''\cos\theta'')^2+(\rho'\sin\theta'-\rho''\sin\theta'')^2}{2z'}} \\
&= \frac{ike^{ikz'}e^{\frac{ik\rho'^2}{2z'}}u_0}{2\pi z'}\int\limits_0^a \mathrm{d}\rho''\ \rho'' e^{\frac{ik\rho''^2}{2z'}}\int\limits_{-\pi}^{\pi}\mathrm{d}\theta''\ e^{\frac{-2ik\rho'\rho''\cos(\theta''-\theta')}{2z'}} \\
&= \frac{ike^{ikz'}e^{\frac{ik\rho'^2}{2z'}}u_0}{z'}\int\limits_0^a \mathrm{d}\rho''\ \rho'' e^{\frac{ik\rho''^2}{2z'}} J_0\left(\frac{k\rho'\rho''}{z'}\right) \\
&= \frac{ika^2u_0}{z'}e^{ikz'}e^{\frac{ik\rho'^2}{2z'}}\int\limits_0^1 \mathrm{d}s\ se^{\frac{ika^2s^2}{2z'}} J_0\left(\frac{ka\rho's}{z'}\right)\ .
\end{aligned} \tag{5.109}$$

In the Fraunhofer (5.60) diffraction integral the terms $e^{ika^2s^2/2z'}$ is disregarded, and consequently $u_1(\mathbf{r}')$ becomes

$$\begin{aligned}
u_1(\mathbf{r}') &= \frac{ika^2u_0}{z'}e^{ikz'}e^{\frac{ik\rho'^2}{2z'}}\int\limits_0^1 \mathrm{d}s\ sJ_0\left(\frac{ka\rho's}{z'}\right) \\
&= \frac{iau_0}{\rho'}e^{ikz'}e^{\frac{ik\rho'^2}{2z'}}J_1\left(\frac{ka\rho'}{z'}\right)\ .
\end{aligned} \tag{5.110}$$

5. The Fresnel diffraction integral (5.56) is given for the present case by

$$\begin{aligned}
u_1(\mathbf{r}') &= \frac{ike^{ikz'}u_0}{2\pi z'}\int\limits_{-a}^{a}\mathrm{d}x''\ e^{\frac{ik(x'-x'')^2}{2z'}}\int\limits_{-b}^{b}\mathrm{d}y''\ e^{\frac{ik(y'-y'')^2}{2z'}} \\
&= \frac{ie^{ikz'}u_0}{2\pi}\int\limits_{\alpha_-}^{\alpha_+}\mathrm{d}X\ e^{\frac{i\pi X^2}{2}}\int\limits_{\beta_-}^{\beta_+}\mathrm{d}X\ e^{\frac{i\pi Y^2}{2}} \\
&= \frac{ie^{ikz'}u_0}{2\pi}\left[F^*(\alpha_+)-F^*(\alpha_-)\right]\left[F^*(\beta_+)-F^*(\beta_-)\right]\ ,
\end{aligned} \tag{5.111}$$

where the Fresnel function $F(q)$ is given by

$$F(q) = \int_0^q \mathrm{d}s\ e^{-\frac{i\pi s^2}{2}}\ , \tag{5.112}$$

and where

$$\alpha_{\pm} = \sqrt{\frac{k}{\pi z'}} \left(\pm a - x'\right) , \tag{5.113}$$

$$\beta_{\pm} = \sqrt{\frac{k}{\pi z'}} \left(\pm b - x'\right) . \tag{5.114}$$

The Fraunhofer diffraction integral (5.60) is given for the present case by

$$\begin{aligned} u_1\left(\mathbf{r}'\right) &= \frac{iku_0}{2\pi z'} e^{ik\left(z' + \frac{x'^2 + y'^2}{2z'}\right)} \int_{-a}^{a} \mathrm{d}x'' e^{-i\frac{kx'}{z'}x''} \int_{-b}^{b} \mathrm{d}y''\ e^{-i\frac{ky'}{z'}y''} \\ &= \frac{2iku_0}{\pi z'} e^{ik\left(z' + \frac{x'^2 + y'^2}{2z'}\right)} ab\,\mathrm{sinc}\left(\frac{kax'}{z'}\right) \mathrm{sinc}\left(\frac{kby'}{z'}\right) . \end{aligned} \tag{5.115}$$

6. With the help of Eqs. (5.56) and (5.46) one obtains

$$\begin{aligned} u_1\left(\mathbf{r}'\right) &= \frac{iku_0 e^{ikz'}}{2\pi z'} \int_{-\infty}^{\infty} \mathrm{d}x''\ \frac{1 + \eta \frac{e^{i\frac{2\pi x''}{a}} + e^{-i\frac{2\pi x''}{a}}}{2}}{2} e^{ik\frac{(x'-x'')^2}{2z'}} \\ &\quad \times \int_{-\infty}^{\infty} \mathrm{d}y''\ e^{ik\frac{(y'-y'')^2}{2z'}} \\ &= -\frac{u_0 e^{ikz'}}{2} \left(1 + \eta e^{-i\frac{2\pi^2 z'}{ka^2}} \cos\frac{2\pi x'}{a}\right) . \end{aligned} \tag{5.116}$$

Consider the case where the distance z' between the aperture and the screen is chosen such that the condition

$$e^{-i\frac{2\pi^2 z'}{ka^2}} = 1 \tag{5.117}$$

is satisfied. For this case one finds with the help of Eq. (5.116) that $\left|u_1\left(\mathbf{r}'\right)\right|^2 = \left|u_0 t\left(x'', y''\right)\right|^2$, 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 $\mathbf{a}_1$ and $\mathbf{a}_2$, which are taken to be given by

$$\mathbf{a}_1 = L\sqrt{3}\hat{\mathbf{y}} , \tag{5.118}$$

$$\mathbf{a}_2 = L\left(\frac{3}{2}\hat{\mathbf{x}} + \frac{\sqrt{3}}{2}\hat{\mathbf{y}}\right) , \tag{5.119}$$

and the so-called basis vectors $\mathbf{c}_1$ and $\mathbf{c}_2$, which are given by

$$\mathbf{c}_1 = 0 , \tag{5.120}$$

$$\mathbf{c}_2 = L\hat{\mathbf{x}} , \tag{5.121}$$

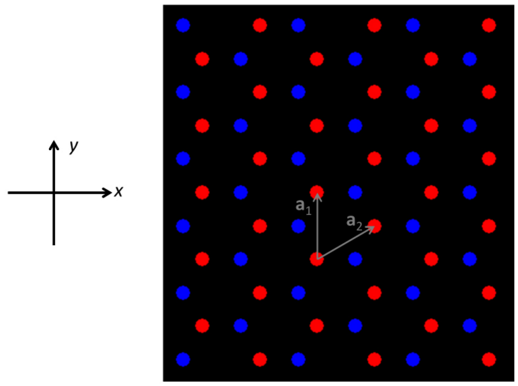

Fig. 5.3. The primitive vectors $\mathbf{a}_1$ and $\mathbf{a}_2$.

in order to specify the locations of the centers of the holes in the array, which are given by

$$\boldsymbol{\rho}_{n_1,n_2,m} = n_1\mathbf{a}_1 + n_2\mathbf{a}_2 + \mathbf{c}_m \ , \tag{5.122}$$

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. 5.3.In the limit $R/L \rightarrow 0$ the holes are treated as point sources. The amplitude in the plane $z = 0^+$ is thus given by

$$u\left(x'', y'', z = 0^+\right) = u_0 \pi R^2 \sum_{n_1,n_2} \sum_{m=1,2} \delta\left(\boldsymbol{\rho}'' - \boldsymbol{\rho}_{n_1,n_2,m}\right) \ , \tag{5.123}$$

where $\boldsymbol{\rho}'' = (x'', y'')$. In the Fraunhofer approximation (5.60) the intensity on a screen is given by

$$I\left(x', y', z_0\right) = \left(\frac{\pi R^2}{\lambda z'}\right)^2 |u_0|^2 \left|\mathcal{A}\left(\frac{x'}{\lambda z_0}, \frac{y'}{\lambda z_0}\right)\right|^2 \ , \tag{5.124}$$

where

$$\mathcal{A}\left(K_x, K_y\right) = \sum_{n_1,n_2} \sum_{m=1,2} e^{-2\pi i \mathbf{K}\cdot\boldsymbol{\rho}_{n_1,n_2,m}} \tag{5.125}$$

and where $\mathbf{K} = K_x\hat{\mathbf{x}} + K_y\hat{\mathbf{y}}$. Consider the so-called reciprocal lattice vectors $\mathbf{b}_1$ and $\mathbf{b}_2$, which are given by

$$\mathbf{b}_1 = \frac{\mathbf{a}_2 \times \hat{\mathbf{z}}}{\mathbf{a}_1 \cdot (\mathbf{a}_2 \times \hat{\mathbf{z}})} = \frac{-\frac{1}{3}\hat{\mathbf{x}} + \frac{1}{\sqrt{3}}\hat{\mathbf{y}}}{L} \ , \tag{5.126}$$

$$\mathbf{b}_2 = \frac{\hat{\mathbf{z}} \times \mathbf{a}_1}{\mathbf{a}_1 \cdot (\mathbf{a}_2 \times \hat{\mathbf{z}})} = \frac{2\hat{\mathbf{x}}}{3L} \ , \tag{5.127}$$

and which satisfy the following relation

$$\mathbf{a}_n \cdot \mathbf{b}_m = \delta_{n,m} \; , \tag{5.128}$$

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

$$\mathbf{K} = K_1 \mathbf{b}_1 + K_2 \mathbf{b}_2 \; , \tag{5.129}$$

one obtains [see Eqs. (5.122), (5.125) and (5.128)]

$$\begin{aligned} \mathcal{A}(K_x, K_y) &= \sum_{n_1, n_2} \sum_{m=1,2} e^{-2\pi i (K_1 \mathbf{b}_1 + K_2 \mathbf{b}_2) \cdot (n_1 \mathbf{a}_1 + n_2 \mathbf{a}_2 + \mathbf{c}_m)} \\ &= T(\mathbf{K}) \sum_{n_1} e^{-2\pi i K_1 n_1} \sum_{n_2} e^{-2\pi i K_2 n_2} \; . \end{aligned} \tag{5.130}$$

where [see Eqs. (5.120) and (5.121)]

$$T(\mathbf{K}) = \sum_{m=1,2} e^{-2\pi i \mathbf{K} \cdot \mathbf{c}_m} = 1 + e^{-2\pi i L \mathbf{K} \cdot \hat{\mathbf{x}}} \; ,$$

or [see Eqs. (5.126) and (5.127)]

$$T(\mathbf{K}) = 1 + e^{\frac{2\pi i}{3}(K_1 - 2K_2)} \; , \tag{5.131}$$

and thus

$$|T(\mathbf{K})|^2 = 4 \cos^2 \left(\frac{\pi (K_1 - 2K_2)}{3} \right) \; . \tag{5.132}$$

For an array of $N \times N$ periods one finds with the help of the identity

$$\sum_{n=-\frac{N}{2}}^{\frac{N}{2}} e^{-i\theta n} = \frac{\sin\left(\frac{N+1}{2}\theta\right)}{\sin\left(\frac{\theta}{2}\right)} \; , \tag{5.133}$$

that

$$\mathcal{A}(K_x, K_y) = T(\mathbf{K}) \, S(K_1) \, S(K_2) \; , \tag{5.134}$$

where

$$S(K) = \frac{\sin((N+1)\pi K)}{\sin(\pi K)} \; . \tag{5.135}$$

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 $\boldsymbol{\rho}' \equiv (x', y') = \boldsymbol{\rho}_{K_1, K_2}$, where

Fig. 5.4. Intensity on the screen for an aperture made of a honeycomb array of holes.

$$\boldsymbol{\rho}_{K_1,K_2} = \lambda z_0 \left(K_1 \mathbf{b}_1 + K_2 \mathbf{b}_2 \right) \ , \tag{5.136}$$

and both K_1 and K_2 are integers. With the help of Eq. (5.132) one finds that for integer values of both K_1 and K_2 the term $|T(\mathbf{K})|^2$ is given by

$$|T(\mathbf{K})|^2 = \begin{cases} 4 \text{ if } \bmod (K_1 - 2K_2, 3) = 0 \\ 1 \qquad\qquad \text{else} \end{cases} \ . \tag{5.137}$$

The intensity on the screen is depicted by Fig. 5.4. The larger spots indicate the points $\boldsymbol{\rho}_{K_1,K_2}$ for which $\bmod (K_1 - 2K_2, 3) = 0$ and $|T(\mathbf{K})|^2 = 4$, whereas $|T(\mathbf{K})|^2 = 1$ for the other points.

8. Consider the function

$$g(s) = \frac{1}{2} \left[1 + \operatorname{sgn} (\cos (s)) \right] \ . \tag{5.138}$$

The following holds $g(s + 2\pi) = g(s)$, i.e. $g(s)$ is periodic, and thus it can be Fourier expanded as

$$
\begin{aligned}
g(s) &= \sum_{m=-\infty}^{\infty} \left(\frac{1}{2\pi} \int_{-\pi}^{\pi} \mathrm{d}s' \; g(s') \, e^{-ims'} \right) e^{ims} \\
&= \sum_{m=-\infty}^{\infty} \left(\frac{1}{2\pi} \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \mathrm{d}s' \; e^{-ims'} \right) e^{ims} \\
&= \sum_{m=-\infty}^{\infty} \frac{1}{\pi m} \sin \frac{\pi m}{2} e^{ims} \; .
\end{aligned}
\tag{5.139}
$$

With the help of the above result (5.139) one finds that the amplitude transmission function (5.87) can be expanded as

$$
\begin{aligned}
t(x'', y'') &= \sum_{m=-\infty}^{\infty} \frac{1}{\pi m} \sin \frac{\pi m}{2} e^{im\gamma r''^2} \\
&= \sum_{m=-\infty}^{\infty} I_m^{1/2} e^{-i \frac{\pi r''^2}{\lambda f_m}} \; ,
\end{aligned}
\tag{5.140}
$$

where f_m, which is given by

$$
f_m = -\frac{\pi}{m\lambda\gamma} \; , \tag{5.141}
$$

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

$$
I_m = \frac{\sin^2 \frac{\pi m}{2}}{\pi^2 m^2} = \begin{cases} \frac{1}{4} & m = 0 \\ \frac{1}{(\pi m)^2} & m \text{ odd} \\ 0 & m \text{ even} \end{cases} \; , \tag{5.142}
$$

represents relative power that is delivered to the m'th focal point, and where

$$
r''^2 = x''^2 + y''^2 \; . \tag{5.143}
$$

Note that for negative values of m the aperture acts as a diverging lens.

9. The Fraunhofer diffraction integral (5.60) for this case is

$$
u_1(\mathbf{r}') = \frac{ik\Phi}{2\pi z'} \int_{-\frac{y''}{\sqrt{3}}}^{\frac{y''}{\sqrt{3}}} \mathrm{d}x'' \int_{0}^{\frac{\sqrt{3}}{2}L} \mathrm{d}y'' \; e^{-i\left(\kappa_x x'' + \kappa_y y''\right)} \; , \tag{5.144}
$$

where

$$
\Phi = e^{ik\left(z' + \frac{x'^2 + y'^2}{2z'}\right)} \; , \tag{5.145}
$$

$k = \omega/c$ and where

$$\kappa_x = \frac{kx'}{z'} \; , \kappa_y = \frac{ky'}{z'} \; , \tag{5.146}$$

thus

$$\begin{aligned} u_1(\mathbf{r}') &= \frac{ik\Phi}{2\pi z'} \int\limits_0^{\frac{\sqrt{3}}{2}L} \mathrm{d}y'' \; e^{-i\kappa_y y''} \int\limits_{-\frac{y''}{\sqrt{3}}}^{\frac{y''}{\sqrt{3}}} \mathrm{d}x'' \; e^{-i\kappa_x x''} \\ &= \frac{i\sqrt{3}k\Phi}{4\pi\kappa_x z'} \left(\frac{e^{-i\left(-\frac{\kappa_x}{2}+\frac{\sqrt{3}\kappa_y}{2}\right)L} - 1}{-\frac{\kappa_x}{2}+\frac{\sqrt{3}\kappa_y}{2}} - \frac{e^{-i\left(\frac{\kappa_x}{2}+\frac{\sqrt{3}\kappa_y}{2}\right)L} - 1}{\frac{\kappa_x}{2}+\frac{\sqrt{3}\kappa_y}{2}} \right) , \end{aligned} \tag{5.147}$$

or

$$u_1(\mathbf{r}') = \frac{i\sqrt{3}kL\Phi}{4\pi\kappa_x z'} \left(\frac{e^{-i\bar{\kappa}\cdot\bar{b}} - 1}{\bar{\kappa}\cdot\bar{b}} - \frac{e^{-i\bar{\kappa}\cdot\bar{a}} - 1}{\bar{\kappa}\cdot\bar{a}} \right) , \tag{5.148}$$

or

$$u_1(\mathbf{r}') = \frac{\sqrt{3}kL^2\Phi}{4\pi z'} \frac{e^{-\frac{i\bar{\kappa}\cdot\bar{b}}{2}} \operatorname{sinc}\frac{\bar{\kappa}\cdot\bar{b}}{2} - e^{-\frac{i\bar{\kappa}\cdot\bar{a}}{2}} \operatorname{sinc}\frac{\bar{\kappa}\cdot\bar{a}}{2}}{\kappa_x L} , \tag{5.149}$$

where

$$\bar{\kappa} = (\kappa_x, \kappa_y) \; , \tag{5.150}$$

and where

$$\operatorname{sinc} q = \frac{\sin q}{q} \; . \tag{5.151}$$

The normalized intensity $|u_1|^2$ is plotted in Fig. 5.5.

10. In general, the Fourier expansion of a periodic function $f(s+2\pi) = f(s)$ is given by

$$f(s) = \frac{a_0}{2} + \sum_{n=1}^{\infty} a_n \cos(ns) + \sum_{n=1}^{\infty} b_n \sin(ns) \; , \tag{5.152}$$

where

$$a_0 = \frac{1}{\pi}\int_{-\pi}^{\pi} \mathrm{d}s \; f(s) \; , \tag{5.153}$$

$$a_n = \frac{1}{\pi}\int_{-\pi}^{\pi} \mathrm{d}s \; f(s)\cos(ns) \; , \tag{5.154}$$

$$b_n = \frac{1}{\pi}\int_{-\pi}^{\pi} \mathrm{d}s \; f(s)\sin(ns) \; . \tag{5.155}$$

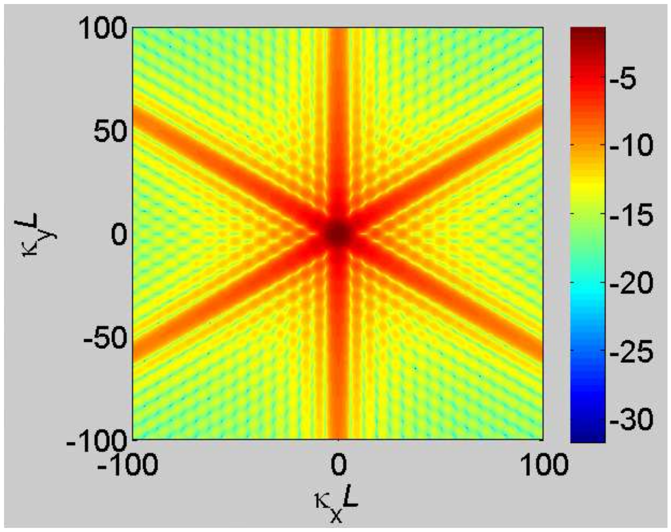

Fig. 5.5. Fraunhofer diffraction of an isosceles triangle aperture. The color coded plot exhibits the normalized intensity (5.148) in a logaritmic scale $\log\left(\left(4\pi z'/\sqrt{3}kL^2\right)^2 |u_1(\kappa_x,\kappa_y)|^2\right)$.

a) The Fourier expansion of the function $\left|\cos\frac{s}{2}\right|$ is given by

$$\left|\cos\frac{s}{2}\right| = \frac{a_0}{2} + \sum_{n=1}^{\infty} a_n \cos(ns) \ , \tag{5.156}$$

where

$$\begin{aligned} a_n &= \frac{1}{\pi}\int_{-\pi}^{\pi} \mathrm{d}s \ \left|\cos\frac{s}{2}\right| \cos(ns) \\ &= \frac{4}{\pi}\frac{\cos(\pi n)}{1-4n^2} \ , \end{aligned} \tag{5.157}$$

and thus the efficiency of the first order is

$$|a_1|^2 = \left(\frac{4}{3\pi}\right)^2 . \tag{5.158}$$

b) For this case Eq. (5.152) yields the expansion

$$e^{\frac{i\phi_0 s}{2\pi}} = \frac{2\sin\frac{\phi_0}{2}}{\frac{\phi_0}{2}} + \sum_n \frac{2\sin\left(\frac{\phi_0}{2}+n\pi\right)}{\frac{\phi_0}{2}+n\pi} e^{ins} \ , \tag{5.159}$$

and thus the efficiency of the first order is $\sin^2(\phi_0/2-\pi)/(\phi_0/2-\pi)^2$.

11. With the help of Eq. (5.115) one finds that the Fraunhofer diffraction integral (5.60) is given by

$$u_1(\mathbf{r}') = \frac{iku_0}{2\pi z'} e^{ik\left(z' + \frac{x'^2+y'^2}{2z'}\right)} \times \left[a_2^2 \operatorname{sinc} \frac{ka_2 x'}{2z'} \operatorname{sinc} \frac{ka_2 y'}{2z'} - a_1^2 \operatorname{sinc} \frac{ka_1 x'}{2z'} \operatorname{sinc} \frac{ka_1 y'}{2z'}\right] . \tag{5.160}$$

References

1. Ref

Index